
Objetivos

El objetivo del presente plan de trabajo es utilizar diversas técnicas de reconocimiento de patrones sobre grandes volúmenes de datos astronómicos. Estas técnicas resultan más atractivas que las tradicionales, las cuales no suelen ser escaleables a gran cantidad de datos. En particular, se utilizarán datos vinculados a galaxias espirales e irregulares cercanas a la Vía Láctea. De esta forma, se pretenden identificar agrupaciones estelares muy jóvenes y obtener sus parámetros fundamentales de una forma sistemática y homogénea.

En base al objetivo planteado, se focalizará la atención sobre los siguientes tópicos que surgen naturalmente de los datos bajo análisis:

- Aplicación de diferentes algoritmos de reconocimiento de patrones y evaluación de su eficiencia haciendo una comparación entre ellos y con los métodos tradicionales
- Identificación y obtención de parámetros de agrupaciones estelares jóvenes extragalácticas
- Evaluación del comportamiento global de dichos parámetros a escala galáctica, prestando atención a la existencia de posibles estructuras jerárquicas en las que eventualmente se puedan diferenciar las asociaciones OB de los cúmulos abiertos
- Búsqueda de correlaciones entre los parámetros obtenidos para las agrupaciones y las características de las galaxias anfitrionas, así como su comparación con aquellos valores típicos aceptados para la Vía Láctea.

Estos estudios apuntan a lograr una mayor comprensión de las propiedades de las poblaciones estelares de galaxias cercanas y de la Vía Láctea. Dado que las diferentes galaxias poseen condiciones ambientales diferentes, como pueden ser su composición química y/o condiciones cinemáticas, el objetivo planteado contribuirá a comprender mejor el desenvolvimiento de los procesos que gobiernan la formación y evolución de las estrellas y de las galaxias en su conjunto. Esta información es muy útil para realizar simulaciones teóricas del funcionamiento y estructura interna de las estrellas. La comparación de dichos modelos con las observaciones mejora nuestro entendimiento de la física básica que domina la naturaleza.

Antecedentes

Las agrupaciones estelares han sido objetos reconocidos desde hace tiempo como laboratorios importantes para la investigación astrofísica. Ellas son muy útiles en varios aspectos, entre los que se pueden destacar los siguientes:

- Los cúmulos estelares contienen muestras estadísticamente significativas de estrellas de aproximadamente de la misma edad con un rango amplio de masas estelares y localizadas en un volumen relativamente pequeño del espacio. Este hecho brinda una forma única para determinar la Función Inicial de Masa (IMF) y verificar si ella manifiesta alguna clase de variación espacial y/o temporal (Scalo 1986; Wilner y Lada 1991; Kroupa et al. 2011).
- Las agrupaciones estelares desempeñan un papel crucial en la comprensión de la estructura de su galaxia anfitriona. Por un lado, los cúmulos abiertos jóvenes y los cúmulos embebidos

se hallan relacionados con la formación estelar reciente. Adicionalmente, estos objetos son importantes trazadores de la estructura espiral del disco galáctico (Baume 2006; Borissova et al. 2011). Por otro lado, los grupos estelares de edad avanzada proporcionan información de la historia temprana de la galaxia. En particular, en lo que respecta a la estructura de la Vía Láctea, varios estudios se han realizado mediante observaciones ópticas e infrarrojas a lo largo del plano galáctico. Por ejemplo, los estudios referidos a los brazos de Scutum-Crux y Carina-Sagitario en el cuarto cuadrante galáctico (p.e. Carraro & Costa 2009; Baume et al. 2009, Molina-Lera et al. 2016 y sus referencias), o la ampliación propuesta para los brazos de Perseo y Norma-Cygnus sobre el tercer cuadrante galáctico (Baume 2006; Baume et al. 2006; Vázquez et al. 2008; Borissova et al. 2011).

- En relación con el proceso de formación estelar, los cúmulos jóvenes permiten esclarecer la forma y las escalas de tiempo en las que estos mecanismos están activos (Sung et al. 1998), así como dependen de los distintos ambientes interestelares de la Vía Láctea o de otras galaxias (Fall & Chandar 2012)

Los trabajos mencionados se han focalizado en mejorar el conocimiento de nuestra propia Galaxia (y de las Nubes de Magallanes, p.e. Baume et al. 2008), pero actualmente hay varios factores que incrementan de forma importante tanto la cantidad de objetos a investigar como la metodología para hacerlo. En este aspecto se pueden destacar los siguientes hitos relevantes.

- En la actualidad existe una gran cantidad de información de las galaxias cercanas (varios Mpc) debido, en gran parte, a que el Telescopio Espacial Hubble (HST) ha permitido obtener datos con alta resolución espacial utilizando varias cámaras de campo amplio (WFPC2; ACS; ver Dalcanton et al. 2009). De esta forma, es posible hacer fotometría precisa multicolor para estudiar estrellas individuales y estructuras en la escala de unos pocos parsecs en dichas galaxias. Consecuentemente, se pueden conducir investigaciones vinculadas con agrupaciones estelares, diferentes poblaciones estelares e historias de formación estelar en ambientes muy diferentes al de la Vía Láctea.
- Existe una enorme cantidad de datos proveniente de las varias observaciones continuas que se están realizando y que se proyectan realizar en modo “survey” (p.e. VVV, <https://vvvsurvey.org/> o LSST; <https://www.lsst.org/>) que necesitan ser estudiados con métodos automáticos. En particular para la identificación y parametrización de nuevas agrupaciones estelares.
- Los algoritmos de “Data Mining” (DM), en particular los relacionados con el reconocimiento automático de patrones, están en la actualidad teniendo una importante revisión y desarrollo (ver p.e. Borne 2009; Ball & Brunner 2010; Schmeja 2011) para su aplicación sobre los datos que surgen de los grandes “surveys”. No obstante, es necesario un cuidadoso estudio de los diferentes métodos y evaluar su eficiencia para obtener resultados relevantes.

Adicionalmente, cabe indicar que el presente plan de trabajo se enmarca dentro de las actividades del Grupo de investigación de Agrupaciones Estelares de la FCAG (ver <http://www.fcaglp.unlp.edu.ar/~gbaume/grupo/> para mas detalles). Este grupo de investigación se ha involucrado en el estudio de agrupaciones estelares desde hace bastante tiempo. En sus estudios se han empleado diversas técnicas observacionales (fotometría, espectroscopía y polarimetría) en diferentes rangos espectrales (óptico, infrarrojo, radio). Estos hechos se reflejan en un importante número de trabajos publicados en revistas internacionales bajo referato (más de 40) de temas relacionados con el propuesto en este plan de trabajo (ver CVs de los integrantes). En particular, durante los últimos años se ha desarrollado el análisis de varias galaxias cercanas con resultados concretos que se han reflejado en la publicación de varios artículos y presentaciones en diferentes reuniones científicas como se detalla a continuación:

- Baume G. & Feinstein C., “Estudio de agrupaciones estelares en NGC 300”, 2009, Boletín de la Asociación Argentina de Astronomía, vol. 52, 97,
- Baume G., Feinstein C., Beletsky Y., Carraro G., Gieles M., Shirmer M., Adamo A., Bastian N. y Davies B., “Estudio en el infrarrojo cercano de agrupaciones estelares en M 83”, 2011, Boletín de la Asociación Argentina de Astronomía, vol. 54 207-210
- Rodríguez M.J., Baume G. & Feinstein C., “Análisis detallado de tres complejos estelares en NGC 300”, 2014, Boletín de la Asociación Argentina de Astronomía, vol. 57 151-153
- Rodríguez M.J., Baume G., Feinstein C., “Búsqueda y análisis de asociaciones estelares en NGC 247”, 2015, Boletín de la Asociación Argentina de Astronomía, vol.58, p.165-167
- Rodríguez M.J., Baume, G., Feinstein, C. “Detailed photometric analysis of young star groups in the galaxy NGC 300”, 2016, Astronomy & Astrophysics, vol 594, id.A34
- Rodríguez M.J., Baume G., Feinstein C., “Systematic study of young star clusters in the galaxy NGC 253, 2017, Multi-Scale Star Formation Meeting, 3 al 7 de Abril, Morelia, Michoacán. México
- Rodríguez M.J., Baume G., Feinstein C., “Identification and analysis of the young population in the starburst galaxy NGC 253”, 2017, enviado para consideración de la revista MNRAS

Actividades y metodología

Como primer paso, se realizará una búsqueda sistemática de agrupaciones estelares en cada galaxia mediante la puesta a punto de códigos desarrollados para esta tarea. Estos programas se basarán en diversos algoritmos utilizados para el reconocimiento automático de patrones. Dichos algoritmos son procedimientos que permiten la agrupación (“clustering”) de un conjunto de vectores en diferentes clases siguiendo un determinado criterio de forma tal que los vectores de un mismo grupo (o “cluster”) comparten propiedades comunes. Esos vectores se encuentran definidos en base a las características de los objetos que representan (estrellas o grupos de estrellas en el presente caso) y los criterios se basan usualmente en una distancia establecida en el espacio de características. Cabe notar que el conocimiento de los grupos puede permitir una descripción sintética de un conjunto de datos multidimensional complejo.

Se nota que este tipo de algoritmos han sido aplicados con éxito en otros campos científicos, y en el caso particular de la astrofísica ellos solo se han utilizado en forma marginal y acotada. En la mayoría de los casos ellos se han basado en analizar sólo las posiciones espaciales para encontrar a los sistemas estelares por sobre-densidades contra el fondo estelar o por un equivalente relacionado con la distribución de distancias entre estrellas. No obstante, estos métodos pueden ampliarse a más dimensiones y utilizar un sistema de búsqueda de agrupaciones donde también se considere la información fotométrica (magnitudes y colores) de los objetos. Resulta obvio por ejemplo que la distribución espacial de las estrellas en una galaxia depende fuertemente de sus brillos y colores, por lo que es un elemento de alta relevancia para discriminar entre los diferentes tipos de poblaciones estelares y entre los objetos que conforman agrupaciones y aquellos pertenecientes al campo.

En particular, se hará énfasis en el estudio y comparación de diversas metodologías como son “partitional vs. hierarchical clustering”, o “statistical vs. nonparametric clustering”. En este sentido, se pueden mencionar los algoritmos “K-means”, “Birch”, “Mean-shift”, “DBSCAN” y “HDBSCAN”. Por otro lado, se explorarán también metodologías de clasificación como “Artificial Neural Networks” (ANN).

Una vez aplicados los diferentes métodos el paso siguiente consistirá en confrontar los resultados obtenidos con cada uno de ellos y con técnicas tradicionales ya establecidas, como pueden ser los métodos de identificación de sobre-densidades a partir de simple conteos (p.e. Lada & Lada 1995) o mediante el criterio de “K Nearest Neighbour” (KNN; Casertano & Hut 1985) y/o métodos similares al denominado “Path Linkage Criterion” (PLC, ver p.e. Battinelli 1991, Pietrzynski

et al. 2001, Rodríguez et al. 2016). Se evaluará entonces la eficiencia de los distintos métodos, y se establecerán sus ventajas y desventajas para ser aplicados sobre los objetos de estudio propuestos.

Adicionalmente, una vez localizadas las agrupaciones y realizada su correlación con aquellas ya conocidas, se efectuarán los siguientes pasos:

- Estudiar detalladamente su distribución y morfología
- Aclarar la naturaleza de estas sobre-densidades (agrupaciones estelares reales o fluctuaciones estadísticas) mediante la construcción y análisis de sus respectivos diagramas color-magnitud y color-color (CMDs y TCDs; ver p.e. Baume et al. 2009)
- En el caso de las agrupaciones reales, construir las respectivas funciones de luminosidad, derivar las correspondientes IMFs y determinar sus parámetros principales (enrojecimiento, edad, metalicidad) comparando los diagramas fotométricos con modelos evolutivos teóricos
- De ser necesario, se realizarán observaciones espectroscópicas complementarias de algunas de las zonas de las agrupaciones más relevantes
- Finalmente, se planea investigar la posible correlación de los parámetros deducidos en los pasos anteriores con la distribución espacial de las agrupaciones estelares dentro de cada galaxia. Ello permitirá deducir el desarrollo de procesos vinculados con la formación y evolución de la galaxia anfitriona como se había planteado en los objetivos del proyecto

Cabe mencionar que tanto la metodología de trabajo como las técnicas a utilizar y desarrollar para abordar la investigación involucran diversas facetas que pueden tener consecuencias positivas en otras ramas de la ciencia o incluso en aplicaciones prácticas. En este sentido se pueden mencionar los siguientes aspectos del proyecto:

- El conocimiento ("know how") sobre el uso y desarrollo de bases datos de gran envergadura y el adecuado procesamiento de imágenes digitales de gran tamaño
- El desarrollo de códigos informáticos especializados para manipular y extraer la información pertinente de las bases de datos e imágenes mencionadas

En relación con la transferencia de los resultados o conocimientos de los estudios, ella se realizará, a nivel profesional, por medio de publicaciones internacionales bajo referato y presentaciones en reuniones científicas. Además, se efectuarán tareas de divulgación a la comunidad.

Factibilidad

Respecto a la factibilidad del plan de trabajo propuesto se pueden mencionar los siguientes puntos:

- Tanto el Director como el co-director del presente plan forman parte de un grupo de trabajo consolidado de la FCAG (UNLP) y del IALP (Conicet – UNLP) orientado, en parte, al estudio de agrupaciones estelares y al medio interestelar. Los integrantes de dicho grupo cuentan con sobrada experiencia en el estudio de agrupaciones estelares en diferentes facetas. Esto es cúmulos abiertos, asociaciones OB y/o cúmulos inmersos. Se destacan además los siguientes puntos:
 - El presente plan representa la continuidad natural de los proyectos de investigación previos y en curso (hasta el presente año) de los que los integrantes forman parte.
 - Existe la financiación asociada por parte de varios proyectos de investigación del

- Conicet y de la UNLP. Estos corresponden concretamente los proyectos: 11/G069 (2003-2005), 11/G084 (2006-2009), 11/G107 (2010-2013); y 11/G130 (2014-2017) de la UNLP, y PIPs 112-200901-00016 y 112-201101-00301 del CONICET.
- Existen vinculaciones con diversos grupos de trabajo en el exterior, en particular la Universidad de Valparaíso, Universidad Nacional de Chile y el European Southern Observatory (en Chile). Además, de integrar de forma activa el grupo de investigadores del proyecto internacional VVV (VISTA Variables in the Vía Láctea; <https://vvvsurvey.org/>)
 - La FCAG (UNLP) y el IALP (Conicet – UNLP) disponen facilidades para llevar a cabo los objetivos planteados, ya que se existe:
 - Acceso a las bases de datos por medio del enlace a Internet de la UNLP
 - Acceso a la bibliografía astronómica necesarias, ya que la biblioteca de la FCAG se halla completamente equipada y con acceso disponible a las publicaciones necesarias
 - Tanto la FCAG como el IALP poseen las facilidades computacionales adecuadas para los procesos de reducción y de análisis de datos correspondientes a los objetivos planteados por medio de equipos individuales y/o clusters de computadoras
 - La postulante que llevará a cabo el presente proyecto posee una preparación académica vinculada con la aplicación de las técnicas observacionales que se utilizarán en este proyecto. Además, ella ha realizado su Tesis de Licenciatura y se encuentra finalizando su Tesis Doctoral orientadas en la presente temática, por lo que se encuentra totalmente capacitada para encarar la presente tarea, la que sería una extensión natural de su actual línea de trabajo.
 - Cualquier otro requerimiento que eventualmente sea necesario para concretar este proyecto, es posible resolverlo mediante la colaboración existente con grupos de investigación del exterior.

Referencias Bibliográficas

- Ball N.M. & Brunner R.J., 2010, International Journal of Modern Physics 19,7
- Battinelli, P., 1991, A&A,244,69
- Baume G., 2006, BAAA 49, 200
- Baume G., Moitinho A., Vázquez R.A., et al., 2006, MNRAS 367, 1441
- Baume G., Noël N.E.D., Costa E. et al. 2008 MNRAS 390, 1683
- Baume G., Carraro G. & Momany Y., 2009, MNRAS 398, 221
- Borissova J., Bonatto C., Kurtev R. et al., 2011, A&A 532,131
- Borne K., 2009, arXiv:0911.0505
- Carraro G. & Costa E., 2009, A&A 493, 71
- Casertano S., & Hut P., 1985, ApJ, 298, 80
- Dalcanton J.J., Williams, B.F., Seth, A.C. et al., 2009, ApJS 183, 67
- Fall S.M. & Chandar R., 2012, ApJ 752, 96
- Kroupa P., Weidner C., Pflamm-Altenburg J. et al., 2011, arXiv1112.3340
- Lada E. A., & Lada C. J., 1995, AJ, 109, 1682
- Molina-Lera J.A., Baume G. Gamen R. et al., 2016, A&A 592, 149
- Pietrzynski G., Gieren W., Fouque P. & Pont F., 2001, A&A 371, 497
- Rodríguez,M.J., Baume, G., Feinstein, C., 2016, A&A 594, 34
- Scalo J.M. 1986, In: Fundamentals of Cosmic Physics, vol. 11, p. 1-278
- Schmeja S., 2011, Astron. Nachr. 332,172
- Sung H., Bessell M.S., Lee S-W., 1998, AJ 115, 734
- Vázquez R.A., May J., Carraro G., Bronfman L., Moitinho A. & Baume G. 2008, ApJ 672, 930
- Wilner D.J. & Lada C.J. 1991, AJ 102, 1050