

Compte-rendu de la réunion gsc4sim de 04 mai 2021

Présents : Anaïs, Catherine, Cécile, Emmanuelle, Esther, Marianne, Mihaela, Pierre F

Ordre du jour : nos scripts de recherche des données dans la sortie gsc4sim, difficultés rencontrées pour écrire ce script, ce qui est faisable par script vs ce qu'on fait à la main, alternatives à la recherche des données dans les grands catalogues - avantages /désavantages programmes find... , vizquery.py, SQL/TAP

État des lieux utilisation gsc4sim pour le traitement des grandes tables :

- Esther : utilisation programmes find en attente d'une procédure commune pour utiliser gsc4sim. Mihaela, Katia, Catherine : utilisation gsc4sim . Catherine utilise quasiment que gsc4sim depuis 1 an, par exemple pour chercher des données Gaia ou SDSS. Mihaela aussi. Catherine nous a montré des exemples de scripts pour extraire des données de gsc4sim , faire le tri des identificateurs (gérer les doublons) .
- malgré les mots clé tel que « !# Look around» qui facilitent la récupération des données, besoin pour reproduire une procédure complète
- besoin commun d'améliorer les scripts qu'on utilise actuellement pour chercher les données gsc4sim
- besoin commun d'un script pour une recherche simultanée dans deux catalogues, exemples Gaia et SDSS. Pour le moment on effectue deux recherches, une sdss et une gaia, et ensuite on ajoute dans le script les données sdss et les données gaia. Si on chèche dans deux catalogues simultanément, la meilleure position va être prise.

Remarques :

1. Gaia DR2 et EDR3 :

Cécile : ne pas mélanger les deux recherches Gaia DR2 et Gaia EDR3 ; on effectue la recherche seulement dans la EDR3 et si rien trouvé, on chèche dans dans DR2. Dans la Gaia DR2 il y a des données surestimées . Par contre, si les auteurs donnent les identificateurs Gaia DR2, nous cherchons les données Gaia DR2 et plus tard il y aura un remplacement global avec la EDR3.

2. **EPIC ou TIC:** compilations ou cross match, pas de données d'origine ; on ne le veut pas dans gsc4sim , vu avec Cécile.

3. **KIC :**les positions sont OK mais les mouvement propres sont souvent fausses, vu avec Cécile. On pourrait ajouter KIC dans gsc4sim, mais on entrera plus beaucoup de KIC.

4. Ellipses d'erreur SDSS :

- pas pour SIMBAD;
- la barre d'erreur des coordonnées SDSS ne doit pas être affichée dans SIMBAD ou dans gsc4sim (en cours MiB et Anaïs)
- coordonnées SDSS en qualité C sans barre d'erreur

Amélioration gsc4sim :

1. Recherche par liste avec gsc4sim :

- ajouter une option pour donner une liste de coordonnées en entrée ;
- notes Anaïs : liste de positions données à asu-tsv ou simplement une boucle awk ; awk vs gawk installés sur les ordinateurs des documentalistes
- Anaïs et Mihaela : ok, testé avec Anaïs en version beta , ça fonctionne. Anaïs a ajouté l'option -f= et mis à jour le help : « -f=: file with only coordinates/names « . L'option -f fonctionne pour une liste de coordonnées et pour une liste de noms. Quelques tests effectués sur des noms SDSS et des

coord Gaia et SDSS. Test comparatif sur ma machine : deux fois plus efficace par liste pour 5000 positions et plus de deux fois plus efficace pour 300 objets :

* recherche dans Gaia DR2, 320 coordonnées : en utilisant la manière actuelle de faire, une commande gsc4sim par ligne ça prend 5min. Si j'utilise la version beta avec l'option -f=coo, ça prend 2min. Note : time bash sdss>sdss_new

* recherche dans sdss12, 5294 coordonnées : avec gsc4sim manière actuelle 66min, avec la version beta 34min et avec findsdss9 : 25min.

- comparaison sortie gsc4sim avec l'option -liste vs ligne par ligne (ligne1 : gsc4sim coo1 ; ligne2 : gsc4sim coo2, etc) pour savoir quels sont les changements à faire dans le script de récupération des données ! Pareil sauf le mot clé « v » absent de la fin du paragraphe dans l'affichage avec l'option -liste

```
buga@cds-doc-mib: ~/Bureau/Simbad_project2020/2021/exemple_gaia_gsc4sim
Fichier Édition Affichage Rechercher Terminal Aide
#---Center: 22.778208 -73.412917 radius=1.5arcsec
#---Source: gaia2
!# Look around: 22.778208 -73.412917 (rs=1.5)...
!# Look around <22.778208 -73.412917>(rs=1.5) in: gaia2
!#--Found possible counterpart at r=0.154arcsec (gaia2)
!From Gaia2
o Gaia DR2 4686266579336706688
a i Gaia DR2 4686266579336706688
c c 01 31 06.8003535734 -73 24 46.510746332 (0) [0.0854 0.0721 90] A 2018yCat.1345...0G
c x 0.0087 [0.0638] A 2018yCat.1345...0G
a dat plx | 1.402 [-0.856 [0.193 0.143 90] A 2018yCat.1345...0G
a dat pm | 1.402 [0.193] | -0.856 [0.143] |ICRS|2018yCat.1345...0G|
c mag G 15.6267 [0.0019] C 2018yCat.1345...0G
#--- gaia2 : 1 matches
#--- 1 total matches
#---Center: 22.778583 -73.433769 radius=1.5arcsec
#---Source: gaia2
!# Look around: 22.778583 -73.433769 (rs=1.5)...
!# Look around <22.778583 -73.433769>(rs=1.5) in: gaia2
!#--Found possible counterpart at r=0.143arcsec (gaia2)
!From Gaia2
o Gaia DR2 4686263525615170688
a i Gaia DR2 4686263525615170688
c c 01 31 06.8898232788 -73 26 01.540628064 (0) [0.0685 0.0560 90] A 2018yCat.1345...0G
c x 0.0087 [0.0638] A 2018yCat.1345...0G
a dat plx | 0.0087 ; 0.0638 | 2018yCat.1345...0G |
c p 0.994 -1.375 [0.158 0.113 90] A 2018yCat.1345...0G
a dat pm | 0.994 [0.158] | -1.375 [0.113] |ICRS|2018yCat.1345...0G|
c mag G 16.6031 [0.0014] C 2018yCat.1345...0G
#--- gaia2 : 1 matches
#--- 1 total matches
v
#---Center: 22.778583 -73.433769 radius=1.5arcsec
#---Source: gaia2

buga@cds-doc-mib: ~/Bureau/gsc4sim
Fichier Édition Affichage Rechercher Terminal Aide
=== Execute beta version ===
### Version du 2021 mai 0 11:58
#---Center: 20.521458 -73.218556 radius=1.5arcsec
#---Source: gaia2
!# Look around: 20.521458 -73.218556 (rs=1.5)...
!# Look around <20.521458 -73.218556>(rs=1.5) in: gaia2
!#--Found possible counterpart at r=0.299arcsec (gaia2)
!From Gaia2
o Gaia DR2 4686495453850419328
a i Gaia DR2 4686495453850419328
c c 01 22 05.1402672919 -73 13 07.080677625 (0) [0.0735 0.0626 90] A 2018yCat.1345...0G
c p 0.788 -1.218 [0.144 0.122 90] A 2018yCat.1345...0G
a dat pm | 0.788 [0.144] | -1.218 [0.122] |ICRS|2018yCat.1345...0G|
c mag G 17.1514 [0.0011] C 2018yCat.1345...0G
#--- gaia2 : 1 matches
#---Center: 20.529708 -73.233361 radius=1.5arcsec
#---Source: gaia2
!# Look around: 20.529708 -73.233361 (rs=1.5)...
!# Look around <20.529708 -73.233361>(rs=1.5) in: gaia2
!#--Found possible counterpart at r=0.061arcsec (gaia2)
!From Gaia2
o Gaia DR2 4686495243393086976
a i Gaia DR2 4686495243393086976
c c 01 22 07.1260642354 -73 14 00.140426628 (0) [0.0636 0.0429 90] A 2018yCat.1345...0G
c x 0.0462 [0.0518] A 2018yCat.1345...0G
a dat plx | 0.0462 ; 0.0518 | 2018yCat.1345...0G |
c p 0.866 -1.264 [0.124 0.083 90] A 2018yCat.1345...0G
a dat pm | 0.866 [0.124] | -1.264 [0.083] |ICRS|2018yCat.1345...0G|
c mag G 16.0061 [0.0011] C 2018yCat.1345...0G
#--- gaia2 : 1 matches
#---Center: 20.542250 -73.281333 radius=1.5arcsec
#---Source: gaia2
```

2. 1ère étape Vizier/cross match pour le tri ; 2ème étape gsc4sim

Voici une alternative pour faire le tri des identificateurs avant gsc4sim, si on préfère le tri ligne par ligne à la place du traitement paragraphe par paragraphe : faire un premier tri en utilisant une recherche dans Vizier ou cross match et ensuite appliquer gsc4sim sur ces objets : ajout id et données .

3. avoir la possibilité de **mettre des conditions / limites sur certains paramètres**, par exemple pour les magnitudes : si on regarde les objets faibles, pouvoir couper à partir d'une certaine valeur . Pour Anais : ajouter une option texte, par ex : Vmag<10 qui est un critère ajouté dans la liste des critères envoyé à Vizier (ex SDSS rmag montré par Cath: 2019RAA.19.127L)

4. option -out=parfile pour afficher les commande en %

5. option pour comparer les magnitudes dans le cas où il y a plusieurs match . Impossible ?!

6. pour les catalogues EPIC, KIC : possibilité d'entrer un nom KIC ou EPIC et d'avoir en retour les coordonnées KIC et respectivement EPIC, pour ensuite pouvoir interroger d'autres catalogues Gaia, SDSS, etc, à partir de ces coordonnées EPIC/KIC. Ainsi nous n'allons pas chercher les données d'origine KIC ou EPIC mais les données d'autres catalogues . Ligne de commentaire à ajouter dans

la sortie gsc4sim , pour signaler que ce ne sont pas les données d'origines, mais les données du catalogue cat1. Pour Anais : nom EPIC/KIC/TIC -> trouver les coordonnées dans Vizier et faire une recherche par position avec le -source ou les catalogues par défaut

Suite :

- travail sur nos différents scripts dans le but d'utiliser toutes le/s même/s.

- Katia et Esther ont déjà mis en place ce qu'il faut pour réunir nos scripts de recherche des données via gsc4sim . Voici le lien du salon rocket.chat pour en discuter :

https://rocket.astro.unistra.fr/group/gsc4sim_cosim.

Pour ceux qui préfèrent par mél, il y a le mél de Katia datant de 04 mai pour l'envoi des scripts.

- réunion avec Anais et/ou Gilles et les documentalistes pour écrire un script qui permet d'exploiter la sortie gsc4sim pour le traitement des grandes tables

- autres points : homogénéiser l'affichage des commentaires dans gsc4sim ex: dans la sortie sdss il y a des '!#--' , 'From', '!' , '! SDSS ' , '! Magnitudes' ;