

# Ottimizzazione dei Garanti accademici

Colli Simone<sup>1</sup> and Merenda Saverio Mattia<sup>1</sup>

<sup>1</sup> `simone.colli@studenti.unipr.it`

<sup>2</sup> `saveriomattia.merenda@studenti.unipr.it`

December 5, 2024

(MM: TODO list (da rimuovere piu' avanti))

(MM:

- Modificare il nome del progetto su github e nel container docker da `pd-project` a `ottimizzazione-garanti-accademici`

)

## Abstract

Questo lavoro presenta l'analisi e l'implementazione di un sistema automatizzato per l'assegnazione dei garanti ai corsi universitari, in conformità ai requisiti ministeriali. L'obiettivo principale è garantire che ogni corso soddisfi i vincoli minimi di docenza, rispettando le regole di distribuzione tra diverse categorie di docenti e ottimizzando l'uso delle risorse disponibili.

Utilizzando la programmazione logica con Answer Set Programming (ASP) (Lifschitz, 2002), abbiamo modellato il problema attraverso fatti, regole e vincoli derivati dai dati ministeriali e universitari. Abbiamo implementato una serie di vincoli per rispettare i minimi richiesti di docenti per corso, evitando sovrapposizioni improprie tra gli incarichi dei docenti e considerando scenari realistici in cui un docente può assumere più ruoli parziali.

L'approccio è stato testato su un dataset reale contenente informazioni su corsi, SSD (Settori Scientifico Disciplinari) e docenti dell'Università degli Studi di Parma. I risultati dimostrano come il sistema possa trovare configurazioni ottimali che soddisfano i requisiti, massimizzando l'efficienza e mantenendo flessibilità nell'assegnazione dei docenti.

## 1 Introduzione

L'assegnazione dei garanti nei corsi universitari costituisce una questione fondamentale per la gestione ottimale delle risorse accademiche. Nell'ambito universitario, il *garante* è un docente responsabile di rappresentare e tutelare la qualità didattica di un corso, garantendo il rispetto dei requisiti disciplinari e istituzionali. I garanti possono appartenere a diverse categorie contrattuali: docenti a tempo indeterminato, docenti a tempo determinato e, in casi eccezionali, docenti a contratto.

La sfida principale consiste nel soddisfare i vincoli ministeriali relativi ai garanti, garantendo al contempo un'allocazione equilibrata e sostenibile delle risorse. Ogni corso deve essere supportato da un numero minimo di garanti, suddivisi tra le diverse fasce contrattuali, per assicurare un livello adeguato di competenza e rappresentatività. Inoltre, è indispensabile che almeno il 50% dei garanti afferisca al Settore Scientifico Disciplinare (SSD) caratterizzante del corso, al fine di garantire la coerenza tra l'offerta formativa e le competenze disciplinari.

Un'ulteriore complessità è rappresentata dall'impiego di docenti a contratto, il cui utilizzo deve essere limitato e subordinato alle sole situazioni in cui non sia possibile soddisfare i requisiti attraverso i docenti strutturati. La necessità di rispettare questi vincoli, combinata con la disponibilità limitata di personale e la necessità di bilanciare il carico di lavoro, rende questo problema una sfida organizzativa e computazionale significativa.

Un primo ostacolo affrontato in questo progetto riguarda la fase di pre-elaborazione dei dati forniti dall'università, descritta nella sezione 2 di questo elaborato. I dati risultano disomogenei e incompleti, richiedendo una significativa pulizia e riorganizzazione prima di poter essere utilizzati efficacemente nel modello ASP. Questa fase ha richiesto lo sviluppo di strumenti dedicati per uniformare e validare i dati.

La sezione 3 esplora il cuore del nostro approccio, ovvero la costruzione del modello ASP. Qui abbiamo implementato regole logiche per soddisfare i vincoli ministeriali e massimizzare l'efficienza dell'assegnazione dei garanti, tenendo conto delle limitazioni sulle fasce contrattuali e sull'impiego dei docenti a contratto.

Successivamente, nella sezione 4, presentiamo i risultati ottenuti su un esempio ridotto, un test che ha permesso di validare il modello in un ambiente controllato e di analizzare la qualità delle soluzioni generate. Infine, la sezione 5 descrive l'applicazione del modello su un dataset completo contenente tutti i corsi universitari, eseguendo un benchmark su larga scala per valutare la capacità del nostro approccio di risolvere problemi reali e complessi.

## 2 Pre processing

(SC: Usa questa macro per scrivere commenti)

I dati analizzati sono stati forniti in formato Excel dall'ufficio xxx(SC: che ufficio?) presentando una struttura tabellare composta da una serie di campi organizzati.

Tuttavia, per consentire un'analisi efficace e priva di errori, è stato necessario effettuare una fase di preprocessing volta a correggere incoerenze, migliorare la qualità e scomporre il dataset in sottoinsiemi più gestibili.

I file forniti presentano al loro interno informazioni relative a:

- Docenti e ricercatori.
- Coperture degli insegnamenti dei vari corsi di laurea.

Dopo un'approfondita analisi di questi file abbiamo deciso di considerare rilevanti per l'analisi solo alcuni campi specifici. Nel file contenente le informazioni relative a docenti e ricercatori abbiamo tenuto in considerazione solo le colonne contenenti la matricola, la fascia e il settore scientifico disciplinare (SSD). Nel file relativo alle coperture abbiamo considerato rilevanti solo le colonne contenenti la matricola del docente, il codice del tipo di corso, il codice del corso di laurea, il settore scientifico disciplinare (SSD).

Per ottenere i dati sopra descritti in una forma coerente e ben definita abbiamo effettuato una serie di operazioni di preprocessing per migliorare la qualità dei dati, così da facilitare l'automatizzazione del workflow generale. Per fare ciò abbiamo utilizzato il linguaggio di programmazione Python, con i moduli pandas, openpyxl.

La maggior parte delle operazioni sono state svolte sul file contenente le coperture degli insegnamenti, suddividendo le attività svolte nei seguenti macroprocessi:

- Gestione delle righe contenenti una o più colonne vuote.
- Scomposizione dei dataset, suddivisione dei dati in sottoinsiemi più piccoli e gestibili.
- Identificazione dei casi particolari in modo univoco ed automatizzato.

Le righe completamente vuote sono state completamente rimosse in quanto non rilevanti per l'analisi, e le righe che presentano le colonne relative al docente (matricola, nome e cognome) vuote sono state isolate all'interno di un file separato in quanto rappresentano degli insegnamenti non assegnati ad alcun docente ('insegnamenti\_senza\_docente.xlsx').

Successivamente per migliorare la gestibilità dei dati, abbiamo creato dei sottoinsiemi minimali e distinti. Il primo passaggio che abbiamo svolto consiste nella suddivisione delle righe che presentano gli insegnamenti tenuti dai docenti/ricercatori, dagli insegnamenti tenuti dai docenti a contratto.

Per identificare i docenti a contratto abbiamo sfruttato il file dei docenti che ci è stato fornito, utilizzato la colonna contenente la matricola per come discriminatore per estrarre e separare, in due file differenti, le righe relative ai docenti a contratto ('docenti\_a\_contratto.xlsx') e quelle relative ai docenti a tempo indeterminato ('coperture.xlsx').

Dopo aver suddiviso i dati come descritto abbiamo proceduto con una piccola modifica dei valori contenuti nella colonna relativa al codice del tipo del corso di laurea, sostituendo tutte le occorrenze di "L" con "LT" per esplicitare i corsi di laurea triennale. Inoltre, abbiamo proceduto con l'identificazione dei casi particolari, utilizzando un controllo incrociato tra due colonne. I casi particolari che abbiamo voluto gestire sono quelli per i corsi di laurea specifici riportati nell'immagine Figure 1.

Per fare ciò abbiamo applicato la seguente logica adattandola al corso da verificare: Tenendo in considerazione la Laurea Triennale di Servizio Sociale abbiamo sostituito il codice del tipo del

Corsi di studio	Docenza di riferim. (minimo)	Professori a tempo indetermin. (minimo)	Ricercatori	Docenti in convenz. art. 6, c. 11, L. 240/2010, oppure docenti art. 1, c. 12, L. 230/2005, oppure docenti a contr. art. 23, L. 240/2010 (massimo)		Docenti di università straniere per CdS interateneo (art. 6, c. 11; art. 23, c. 3, L. 240/2010)
				Tot.	di cui docenti a contratto	
LT	9	5	4	3	2	4
LM	6	4	2	2	1	3
LMCU 5 anni	15	8	7	5	3	7
LMCU 6 anni	18	10	8	6	4	9
LT Servizio Sociale LT Scienze Motorie	5	3	2	2	1	2
LT Prof. sanitarie LT a orient. profess. LM Servizio Sociale LM Scienze Motorie	4	2	2	1	1	2
LM Infermieristica	3	1	2	1	1	1

Figure 1: Casi particolari

corso (precedente "LT") con "LTSS" solo nelle righe dove la colonna relativa alla descrizione del corso conteneva un pattern contenente entrambe le keyword "serviz" "social".

In fine per avere la possibilità di estrarre su richiesta dei piccoli dataset tramite parsing a linea di comando abbiamo utilizzato dei file csv di supporto così da creare un collegamento rapido e leggero tra il codice identificativo del corso e la descrizione del corso, ed il codice del corso e le matricole dei docenti che hanno almeno una cattedra in quel corso di laurea.

### 3 ASP

In the *Introduction*, you should start by providing background and context for your study, highlighting the importance and relevance of the topic. Then, clearly identify the research problem or gap in your study's existing literature, explaining why it is significant. State the research questions or objectives your study aims to answer, ensuring they are directly linked to the identified problem. Justify the need for your research by discussing its potential contributions or impact on the field. You may also include a brief overview of the methodology, especially if it's novel or crucial to your study's contribution. Additionally, define the scope and limitations of your research, clarifying what the study will and will not cover. If applicable, present your main thesis statement or hypothesis. Optionally, you can conclude the introduction with a brief outline of the paper's structure to guide the reader.

### 4 Esempio giocattolo

In the *Introduction*, you should start by providing background and context for your study, highlighting the importance and relevance of the topic. Then, clearly identify the research problem or gap in your study's existing literature, explaining why it is significant. State the research questions or objectives your study aims to answer, ensuring they are directly linked to the identified problem. Justify the need for your research by discussing its potential contributions or impact on the field. You may also include a brief overview of the methodology, especially if it's novel or crucial to your study's contribution. Additionally, define the scope and limitations of your research, clarifying what the study will and will not cover. If applicable, present your main thesis statement or hypothesis. Optionally, you can conclude the introduction with a brief outline of the paper's structure to guide the reader.

### 5 Valutazione sperimentale

In the *Introduction*, you should start by providing background and context for your study, highlighting the importance and relevance of the topic. Then, clearly identify the research problem or gap in your study's existing literature, explaining why it is significant. State the research questions or objectives your study aims to answer, ensuring they are directly linked to the identified problem. Justify the need for your research by discussing its potential contributions or impact on the field. You may also include a brief overview of the methodology, especially if it's novel or crucial to your study's contribution. Additionally, define the scope and limitations of your research,

clarifying what the study will and will not cover. If applicable, present your main thesis statement or hypothesis. Optionally, you can conclude the introduction with a brief outline of the paper's structure to guide the reader.

## 6 Conclusion

A strong *Conclusion* typically encompasses several essential elements. It begins with a restatement of the thesis or main idea, reinforcing the core message of the paper. This is followed by a summary of the key points discussed throughout the text, which helps to consolidate the main arguments. The conclusion should also include final thoughts or reflections on the topic, providing a thoughtful discussion wrap-up. (**fellows'research'2021**)

### *Acknowledgments*

Questo progetto è stato realizzato nel corso di Programmazione Dichiarativa (a.a. 2024-25), presso l'Università degli Studi di Parma.

## References

Lifschitz, Vladimir (2002). “Answer set programming and plan generation”, *Artificial Intelligence*, Vol. 138 No. 1. Knowledge Representation and Logic Programming, pp. 39–54. ISSN: 0004-3702. DOI: [https://doi.org/10.1016/S0004-3702\(02\)00186-8](https://doi.org/10.1016/S0004-3702(02)00186-8). **available at:** <https://www.sciencedirect.com/science/article/pii/S0004370202001868>.