Risultati problemi senza aggiunta azioni

**BLOCKSWORLD**

File: generation\_output\_100p\_NoAction.json / generation\_results\_100p\_NoAction.txt

Correct plans: 2024

Total plans: 2028

Coverage: 99.80%

Total time: 922.52

**SATELLITE**

File: generation\_output\_100p\_NoAction.json / generation\_results\_100p\_NoAction.txt

Correct plans: 460

Total plans: 658

Coverage: 69.91%

Total time: 577.26

**DRIVERLOG**

File: generation\_output\_100p\_NoAction.json / generation\_results\_100p\_NoAction.txt

Correct plans: 701

Total plans: 704

Coverage: 99.57%

Total time: 1045.44

**ZENOTRAVEL**

File: generation\_output\_100p\_NoAction.json / generation\_results\_100p\_NoAction.txt

Correct plans: 670

Total plans: 670

Coverage: 100%

Total time: 732.30

Risultati problemi con l’aggiunta delle azioni

**BLOCKSWORLD**

File: generation\_output\_100p\_actions.json / generation\_results\_100p\_actions.txt

Correct plans: 19761

Total plans: 35118

Coverage: 56.27%

Total time: 20817.59

**SATELLITE**

File: generation\_output\_100p\_actions.json / generation\_results\_100p\_actions.txt

Correct plans: 2612

Total plans: 12954

Coverage: 20.16%

Total time: 7089.86

**DRIVERLOG**

File: generation\_output\_100p\_actions.json / generation\_results\_100p\_actions.txt

Correct plans: 10658

Total plans: 15356

Coverage: 69.41%

Total time: 27490.57

**ZENOTRAVEL**

File: generation\_output\_100p\_actions.json / generation\_results\_100p\_actions.txt

Correct plans: 5278

Total plans: 12788

Coverage: 41.27

Total time: 12087.20

RF e CV delle 3 metriche

**metrica: piano valido e più corto**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Domain | RF | | CV | |
| Mean | STD | Mean | STD |
| Blocksworld | 0.350 | 0.230 | 0.318 | 0.233 |
| Driverlog | 0.596 | 0.225 | 0.499 | 0.264 |
| Satellite | 0.405 | 0.237 | 0.344 | 0.239 |
| Zenotravel | 0.704 | 0.169 | 0.577 | 0.266 |

**metrica: maggior numero di goals soddisfatti**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Domain | RF | | CV | |
| Mean | STD | Mean | STD |
| Blocksworld | 0.276 | 0.232 | 0.258 | 0.231 |
| Driverlog | 0.296 | 0.174 | 0.271 | 0.182 |
| Satellite | 0.575 | 0.174 | 0.512 | 0.178 |
| Zenotravel | 0.549 | 0.178 | 0.515 | 0.175 |

**metrica: reward calcolata come: (numero goal soddisfatti / numero totali dei goals) (- 1)\*** \*se check\_violated\_preconditions

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Domain | RF | | CV | |
| Mean | STD | Mean | STD |
| Blocksworld | 0.293 | 0.180 | 0.265 | 0.184 |
| Driverlog | 0.367 | 0.173 | 0.290 | 0.175 |
| Satellite | 0.531 | 0.187 | 0.381 | 0.220 |
| Zenotravel | 0.579 | 0.128 | 0.436 | 0.174 |

**Accuracy delle 5 metriche**

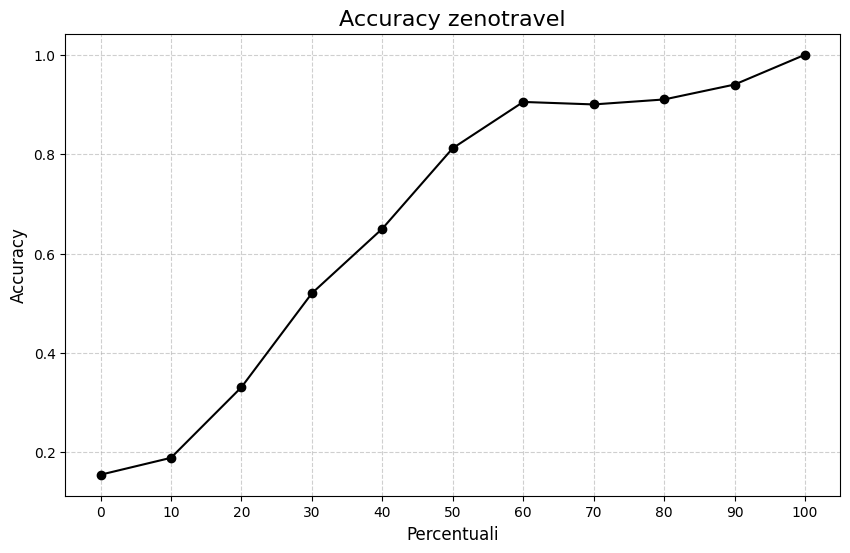
**metrica: piano valido e più corto**

zenotravel: [0.155, 0.189, 0.331, 0.52, 0.65, 0.812, 0.905, 0.9, 0.91, 0.94, 1.0]

drivelog: [0.12, 0.15, 0.202, 0.328, 0.542, 0.648, 0.73, 0.825, 0.9, 0.89, 0.98]

blocksworld: [0.033, 0.031, 0.051, 0.079, 0.246, 0.257, 0.267, 0.435, 0.569, 0.752, 0.87]

satellite: [0.147, 0.201, 0.225, 0.18, 0.214, 0.243, 0.374, 0.445, 0.59, 0.81, 0.995]

 Immagine che contiene testo, linea, Diagramma, diagramma

Il contenuto generato dall'IA potrebbe non essere corretto.

Immagine che contiene linea, testo, Diagramma, diagramma

Il contenuto generato dall'IA potrebbe non essere corretto.Immagine che contiene linea, Diagramma, diagramma, testo

Il contenuto generato dall'IA potrebbe non essere corretto.

**metrica: maggior numero di goals soddisfatti**

zenotravel: [0.139, 0.144, 0.212, 0.283, 0.402, 0.523, 0.661, 0.771, 0.816, 0.91, 0.935]

drivelog: [0.147, 0.15, 0.159, 0.173, 0.215, 0.266, 0.302, 0.378, 0.417, 0.493, 0.536]

blocksworld: [0.04, 0.04, 0.043, 0.072, 0.173, 0.177, 0.237, 0.323, 0.408, 0.577, 0.752]

satellite: [0.132, 0.191, 0.239, 0.28, 0.401, 0.518, 0.677, 0.806, 0.892, 0.963, 0.968]

Immagine che contiene linea, Diagramma, diagramma, testo

Il contenuto generato dall'IA potrebbe non essere corretto. Immagine che contiene testo, Diagramma, linea, diagramma

Il contenuto generato dall'IA potrebbe non essere corretto. Immagine che contiene testo, linea, Diagramma, diagramma

Il contenuto generato dall'IA potrebbe non essere corretto.Immagine che contiene linea, Diagramma, diagramma, testo

Il contenuto generato dall'IA potrebbe non essere corretto.

**metrica: reward calcolata come: (numero goal soddisfatti / numero totali dei goals) (- 1)\*** \*se check\_violated\_preconditions

zenotravel: [0.162, 0.165, 0.258, 0.299, 0.438, 0.665, 0.659, 0.772, 0.889, 0.95, 0.972]

drivelog: [0.148, 0.143, 0.181, 0.194, 0.284, 0.35, 0.384, 0.458, 0.57, 0.606, 0.663]

blocksworld: [0.048, 0.048, 0.048, 0.077, 0.158, 0.158, 0.174, 0.366, 0.466, 0.637, 0.815]

satellite: [0.132, 0.226, 0.233, 0.3, 0.497, 0.494, 0.617, 0.669, 0.692, 0.895, 0.972]

Immagine che contiene diagramma, linea, Diagramma, testo

Il contenuto generato dall'IA potrebbe non essere corretto. Immagine che contiene testo, linea, Diagramma, diagramma

Il contenuto generato dall'IA potrebbe non essere corretto.

Immagine che contiene testo, linea, Diagramma, diagramma

Il contenuto generato dall'IA potrebbe non essere corretto. Immagine che contiene linea, diagramma, Diagramma, testo

Il contenuto generato dall'IA potrebbe non essere corretto.

**Analisi delle soluzioni delle 5 metriche**

**metrica: piano valido e più corto**

* Caso1: piano generato per il goal obiettivo non valido
* Caso2: piano generato per il goal obiettivo valido, ma con lunghezza maggiore degli altri piani validi generati per gli altri goals
* Caso3: piano generato per il goal obiettivo valido, ma con lunghezza minima uguale ad altri paini validi generati per gli altri goals
* Caso4: piano generato per il goal obiettivo valido con lunghezza minima
* Caso5: piani generati per tutti i goals non validi

Immagine che contiene testo, diagramma, schermata, Parallelo

Il contenuto generato dall'IA potrebbe non essere corretto. Immagine che contiene testo, diagramma, schermata, Parallelo

Il contenuto generato dall'IA potrebbe non essere corretto. Immagine che contiene testo, diagramma, schermata, linea

Il contenuto generato dall'IA potrebbe non essere corretto.Immagine che contiene testo, diagramma, schermata, Carattere

Il contenuto generato dall'IA potrebbe non essere corretto.

**metrica: maggior numero di goals soddisfatti**

* Caso1: piano generato per il goal obiettivo con numero di goals soddisfatti minore degli altri piani generati per gli altri goals
* Caso2: piano generato per il goal obiettivo con numero di goals soddisfatti uguale ad altri piani generati per gli altri goals
* Caso3: piano generato per il goal obiettivo con numero goals soddisfatti maggiore degli altri piani generati per gli altri goals

Immagine che contiene testo, diagramma, schermata, Parallelo

Il contenuto generato dall'IA potrebbe non essere corretto. Immagine che contiene testo, diagramma, schermata, Parallelo

Il contenuto generato dall'IA potrebbe non essere corretto.

Immagine che contiene testo, diagramma, schermata, Parallelo

Il contenuto generato dall'IA potrebbe non essere corretto. Immagine che contiene testo, diagramma, schermata, Parallelo

Il contenuto generato dall'IA potrebbe non essere corretto.

**metrica: reward calcolata come: (numero goal soddisfatti / numero totali dei goals) (- 1)\*** \*se check\_violated\_preconditions

* Caso1: piano generato per il goal obiettivo con la reward minore degli altri piani generati per gli altri goals
* Caso2: piano generato per il goal obiettivo con la reward uguale ad altri piani generati per gli altri goals
* Caso3: piano generato per il goal obiettivo con la reward maggiore degli altri piani generati per gli altri goals

Immagine che contiene testo, diagramma, schermata, Parallelo

Il contenuto generato dall'IA potrebbe non essere corretto.Immagine che contiene testo, diagramma, schermata, Parallelo

Il contenuto generato dall'IA potrebbe non essere corretto.

Immagine che contiene testo, diagramma, schermata, Parallelo

Il contenuto generato dall'IA potrebbe non essere corretto. Immagine che contiene testo, diagramma, schermata, Parallelo

Il contenuto generato dall'IA potrebbe non essere corretto.

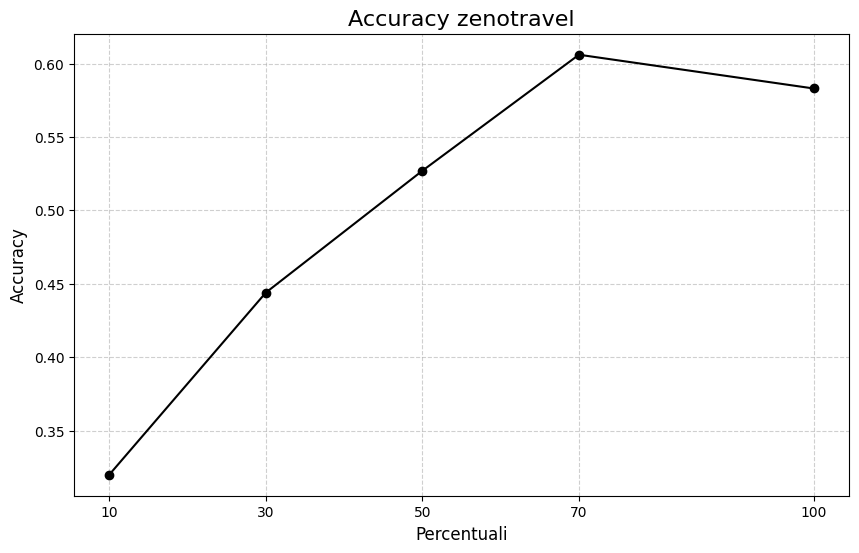
**GR OFFLINE**

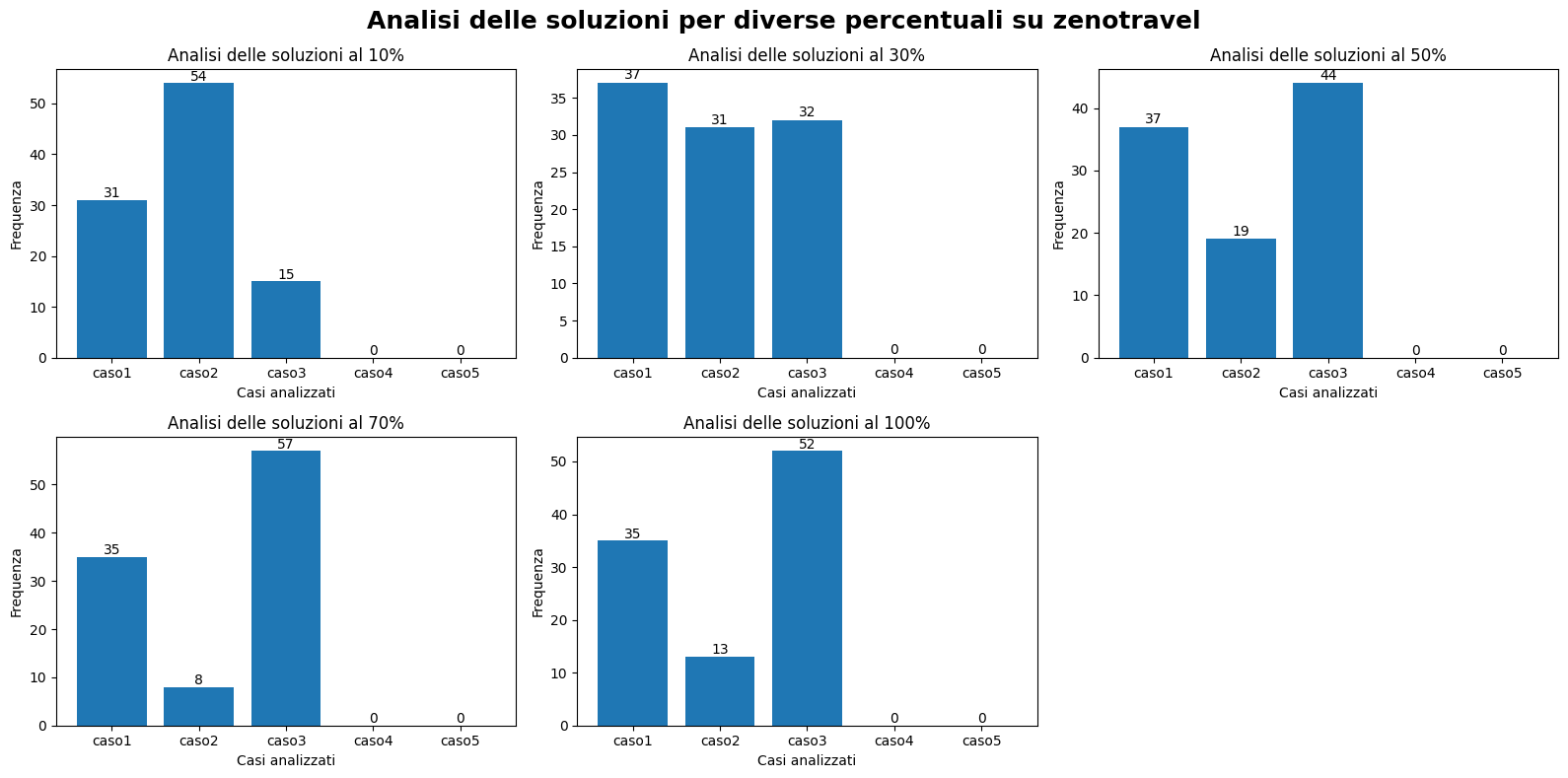
**Metrica: plan distance, ovvero quante azioni osservate compaiono nel piano generato da PlanGPT**

* Caso1: piano generato per il goal obiettivo valido e con la metrica minore degli altri piani generati per gli altri goals
* Caso2: piano generato per il goal obiettivo valido e con la metrica uguale ad altri piani generati per gli altri goals
* Caso3: piano generato per il goal obiettivo valido e con la metrica maggiore degli altri piani generati per gli altri goals
* Caso4: piano generato per il goal obiettivo non valido
* Caso5: i piani generati per tutti i goals risultano essere non validi

**Zenotravel**

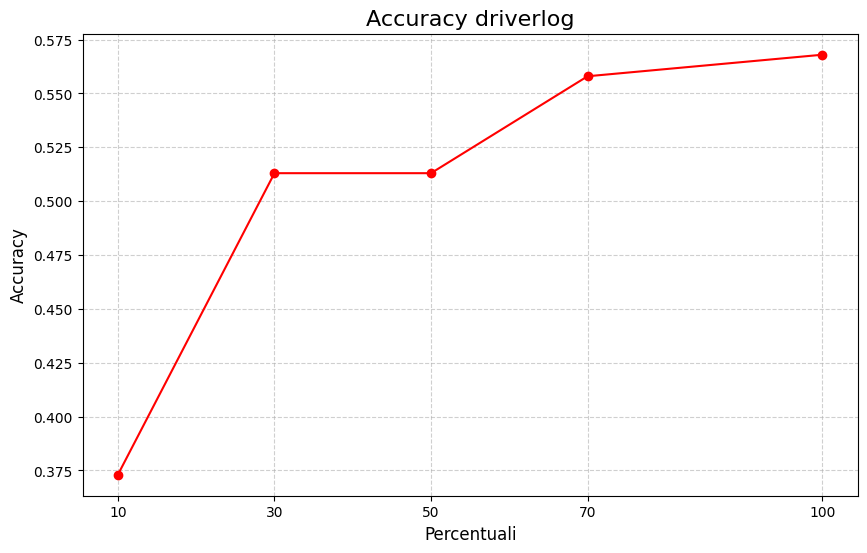
Accuracy: [0.32, 0.444, 0.527, 0.606, 0.583]

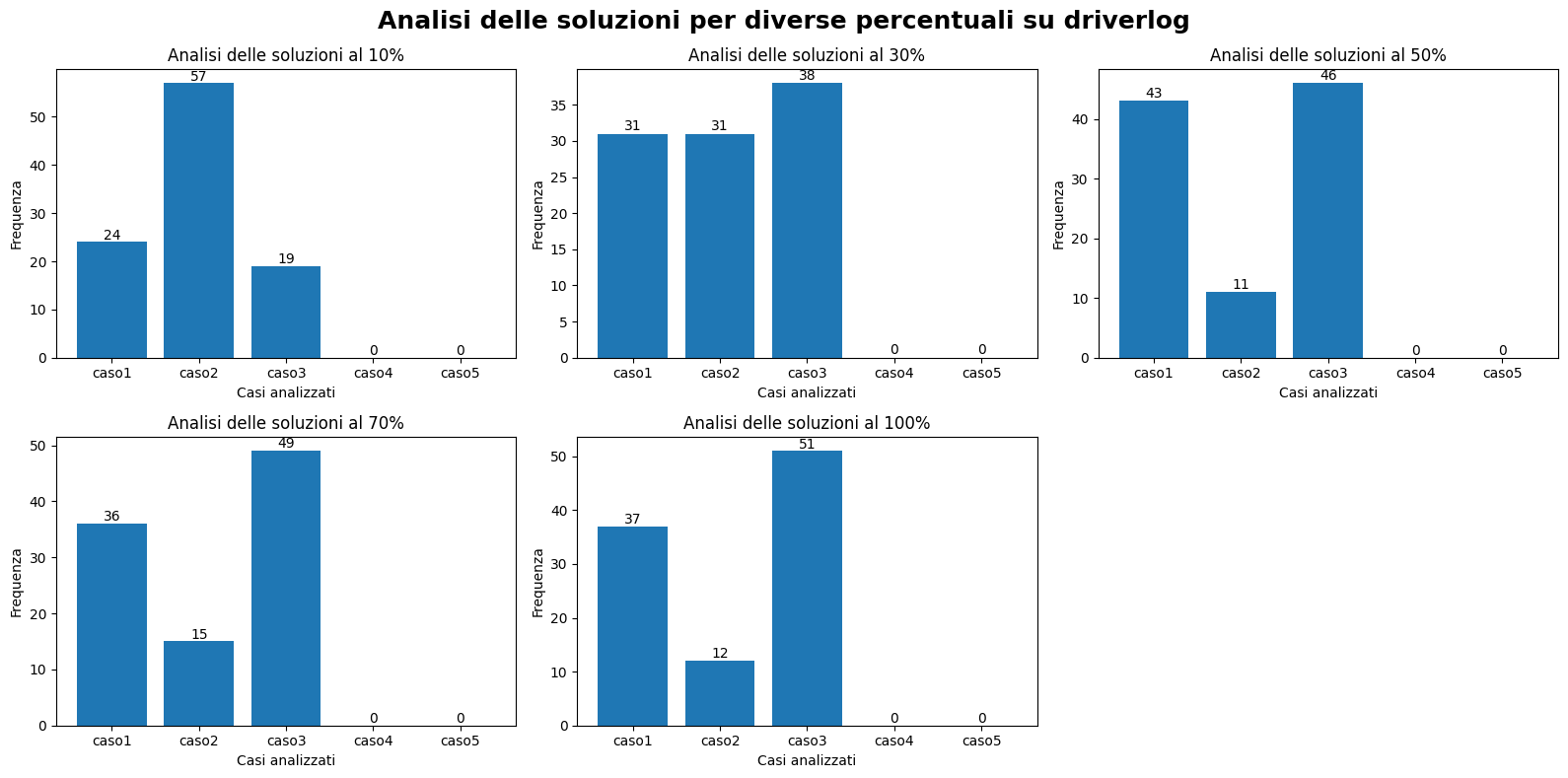




**Driverlog**

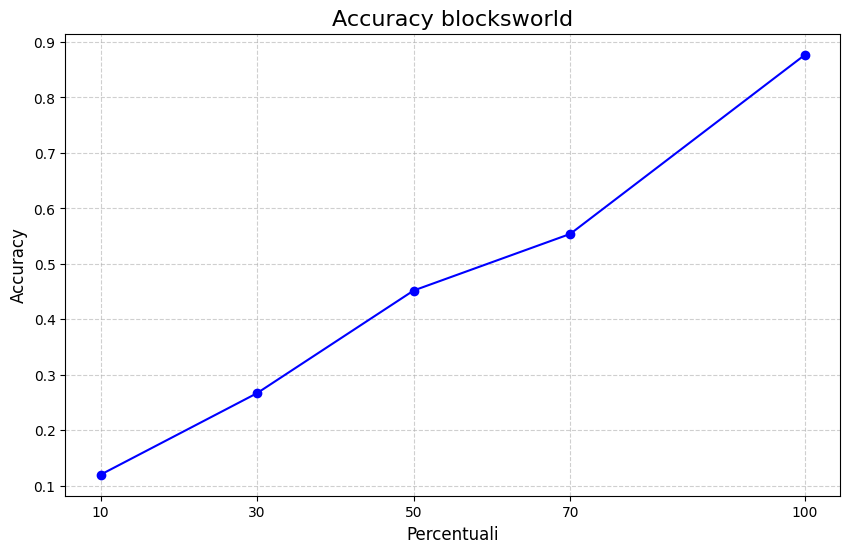
Accuracy: [0.373, 0.513, 0.513, 0.558, 0.568]

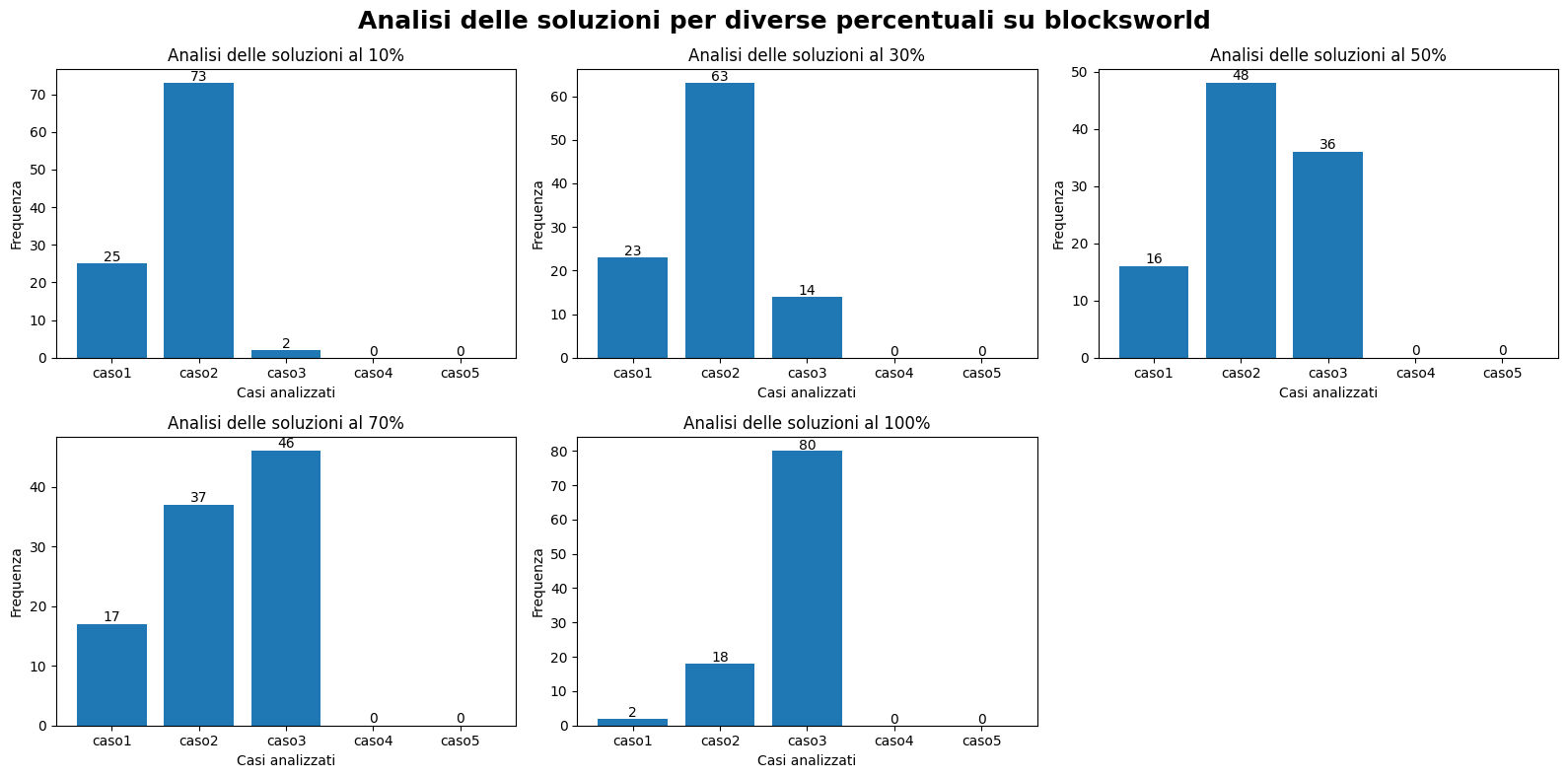


****

**Blocksworld**

Accuracy: [0.12, 0.267, 0.452, 0.554, 0.877]

****

****

**Satellite**

Accuracy: [0.274, 0.359, 0.347, 0.363, 0.376]

Immagine che contiene linea, diagramma, Diagramma, testo

Il contenuto generato dall'IA potrebbe non essere corretto.

**Immagine che contiene testo, diagramma, schermata, Diagramma

Il contenuto generato dall'IA potrebbe non essere corretto.**