Risultati problemi senza aggiunta azioni

**BLOCKSWORLD**

File: generation\_output\_100p\_NewTest\_union.json / generation\_results\_100p\_NewTest\_union.txt

Correct plans: 1843 (errori nei piani generati con il modello che accetta clean nei goals)

Total plans: 2028

Coverage: 90.88%

Total time: 692.095

**SATELLITE**

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

Correct plans: 316 (errori: 56 in problemi di Pereira e 285 in problemi di Lorenzo)

Total plans: 658

Coverage: 48.02%

Total time: 395.2155

**DRIVERLOG**

File: generation\_output\_100p\_NewTest1.json / generation\_results\_100p\_NewTest1.txt

Correct plans: 652

Total plans: 704

Coverage: 92.61%

Total time: 549.575

**ZENOTRAVEL**

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

Correct plans: 653

Total plans: 670

Coverage: 97.46%

Total time: 402.095

Risultati problemi con l’aggiunta delle azioni

**BLOCKSWORLD**

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

Correct plans: 17654

Total plans: 35118

Coverage: 50.27%

Total time: 14047.813

**SATELLITE**

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

Correct plans: 1971

Total plans: 12954

Coverage: 15.21%

Total time: 4556.569

**DRIVERLOG**

File: generation\_output\_100p\_ actions \_NewTest.json / generation\_results\_100p\_ actions\_NewTest.txt

Correct plans: 8046

Total plans: 15356

Coverage: 52.40%

Total time: 13022.27

**ZENOTRAVEL**

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

Correct plans: 4140

Total plans: 12788

Coverage: 32.37%

Total time: 6884.950

RF e CV delle 3 metriche

**metrica: piano valido e più corto**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Domain | RF | | CV | |
| Mean | STD | Mean | STD |
| Blocksworld | 0.321 | 0.230 | 0.292 | 0.236 |
| Driverlog | 0.570 | 0.239 | 0.438 | 0.273 |
| Satellite | 0.376 | 0.241 | 0.331 | 0.236 |
| Zenotravel | 0.647 | 0.206 | 0.532 | 0.274 |

**metrica: maggior numero di goals soddisfatti**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Domain | RF | | CV | |
| Mean | STD | Mean | STD |
| Blocksworld | 0.280 | 0.220 | 0.258 | 0.222 |
| Driverlog | 0.374 | 0.203 | 0.322 | 0.212 |
| Satellite | 0.593 | 0.173 | 0.526 | 0.182 |
| Zenotravel | 0.586 | 0.157 | 0.541 | 0.158 |

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

**metrica: calcola il numero di goal soddisfatti ma penalizzando sulla lunghezza (no\_penality = false)**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Domain | RF | | CV | |
| Mean | STD | Mean | STD |
| Blocksworld | 0.299 | 0.202 | 0.279 | 0.210 |
| Driverlog | 0.374 | 0.180 | 0.321 | 0.193 |
| Satellite | 0.535 | 0.189 | 0.396 | 0.230 |
| Zenotravel | 0.590 | 0.138 | 0.509 | 0.172 |

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Domain | RF | | CV | |
| Mean | STD | Mean | STD |
| Blocksworld | 0.151 | 0.232 | 0.093 | 0.218 |
| Driverlog | 0.176 | 0.303 | 0.121 | 0.268 |
| Satellite | 0.193 | 0.228 | 0.081 | 0.193 |
| Zenotravel | 0.217 | 0.267 | 0.099 | 0.212 |

**metrica: calcola il numero di goal soddisfatti ma penalizzando sulla lunghezza (no\_penality = true)**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Domain | RF | | CV | |
| Mean | STD | Mean | STD |
| Blocksworld | 0.087 | 0.201 | 0.065 | 0.189 |
| Driverlog | 0.164 | 0.307 | 0.126 | 0.275 |
| Satellite | 0.155 | 0.234 | 0.090 | 0.213 |
| Zenotravel | 0.175 | 0.288 | 0.112 | 0.267 |

**Accuracy delle 5 metriche**

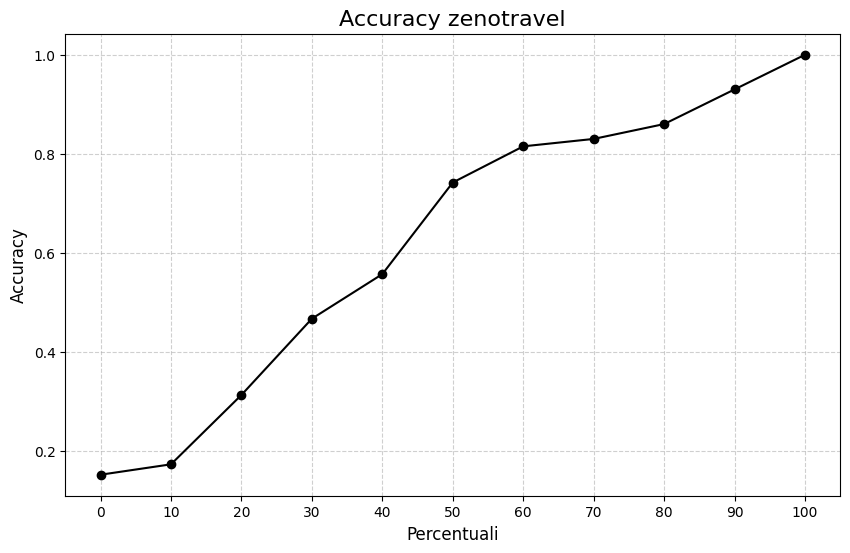
**metrica: piano valido e più corto**

zenotravel: [0.152, 0.173, 0.313, 0.467, 0.557, 0.742, 0.815, 0.83, 0.86, 0.93, 1.0]

drivelog: [0.155, 0.185, 0.2, 0.375, 0.552, 0.585, 0.615, 0.715, 0.83, 0.855, 0.975]

blocksworld: [0.038, 0.043, 0.049, 0.081, 0.188, 0.198, 0.218, 0.351, 0.524, 0.757, 0.833]

satellite: [0.124, 0.166, 0.145, 0.122, 0.182, 0.233, 0.342, 0.425, 0.55, 0.8, 0.995]

 Immagine che contiene testo, linea, Diagramma, diagramma

Descrizione generata automaticamente Immagine che contiene testo, linea, Diagramma, diagramma

Descrizione generata automaticamente Immagine che contiene linea, Diagramma, diagramma, testo

Descrizione generata automaticamente

**metrica: maggior numero di goals soddisfatti**

zenotravel: [0.137, 0.146, 0.197, 0.276, 0.441, 0.589, 0.724, 0.851, 0.902, 0.942, 0.966]

drivelog: [0.158, 0.161, 0.178, 0.202, 0.269, 0.328, 0.4, 0.492, 0.527, 0.59, 0.665]

blocksworld: [0.04, 0.04, 0.045, 0.066, 0.166, 0.165, 0.239, 0.335, 0.417, 0.598, 0.778]

satellite: [0.136, 0.236, 0.226, 0.303, 0.467, 0.523, 0.692, 0.833, 0.912, 0.962, 0.972]

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

Descrizione generata automaticamente** **Immagine che contiene testo, linea, Diagramma, diagramma

Descrizione generata automaticamente** Immagine che contiene testo, linea, Diagramma, diagramma

Descrizione generata automaticamente Immagine che contiene linea, diagramma, Diagramma, testo

Descrizione generata automaticamente

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

zenotravel: [0.151, 0.159, 0.213, 0.283, 0.461, 0.608, 0.721, 0.824, 0.897, 0.952, 0.966]

drivelog: [0.155, 0.159, 0.181, 0.204, 0.275, 0.331, 0.392, 0.484, 0.529, 0.6, 0.669]

blocksworld: [0.051, 0.051, 0.052, 0.084, 0.18, 0.174, 0.199, 0.352, 0.45, 0.665, 0.818]

satellite: [0.151, 0.233, 0.234, 0.32, 0.472, 0.508, 0.615, 0.609, 0.7, 0.892, 0.973]

Immagine che contiene linea, Diagramma, diagramma, testo

Descrizione generata automaticamente Immagine che contiene linea, testo, Diagramma, diagramma

Descrizione generata automaticamente

Immagine che contiene testo, linea, Diagramma, diagramma

Descrizione generata automaticamente Immagine che contiene linea, Diagramma, diagramma, testo

Descrizione generata automaticamente

**metrica: calcola il numero di goal soddisfatti ma penalizzando sulla lunghezza (no\_penality = false)**

zenotravel: [0.15, 0.163, 0.163, 0.184, 0.258, 0.33, 0.255, 0.168, 0.243, 0.201, 0.157]

drivelog: [0.16, 0.16, 0.177, 0.177, 0.21, 0.21, 0.215, 0.175, 0.195, 0.131, 0.05]

blocksworld: [0.085, 0.075, 0.07, 0.078, 0.168, 0.193, 0.186, 0.23, 0.242, 0.246, 0.111]

satellite: [0.158, 0.213, 0.209, 0.191, 0.238, 0.212, 0.223, 0.21, 0.184, 0.202, 0.113]

 Immagine che contiene testo, linea, Diagramma, diagramma

Descrizione generata automaticamente Immagine che contiene linea, diagramma, testo, Diagramma

Descrizione generata automaticamente Immagine che contiene diagramma, linea, Diagramma

Descrizione generata automaticamente

**metrica: calcola il numero di goal soddisfatti ma penalizzando sulla lunghezza (no\_penality = true)**

zenotravel: [0.15, 0.164, 0.187, 0.187, 0.243, 0.213, 0.202, 0.165, 0.178, 0.116, 0.063]

drivelog: [0.16, 0.16, 0.177, 0.177, 0.187, 0.181, 0.194, 0.159, 0.186, 0.13, 0.043]

blocksworld: [0.08, 0.08, 0.092, 0.092, 0.099, 0.102, 0.096, 0.103, 0.088, 0.079, 0.042]

satellite:

Immagine che contiene diagramma, linea, Diagramma

Descrizione generata automaticamente Immagine che contiene testo, Diagramma, linea, diagramma

Descrizione generata automaticamente Immagine che contiene linea, testo, Diagramma, diagramma

Descrizione generata automaticamente Immagine che contiene linea, diagramma, Diagramma

Descrizione generata automaticamente

**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

Descrizione generata automaticamente Immagine che contiene testo, diagramma, schermata, Parallelo

Descrizione generata automaticamente Immagine che contiene testo, diagramma, schermata, Carattere

Descrizione generata automaticamenteImmagine che contiene testo, diagramma, schermata, Carattere

Descrizione generata automaticamente

**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

Descrizione generata automaticamente** **Immagine che contiene testo, diagramma, schermata, Parallelo

Descrizione generata automaticamente**

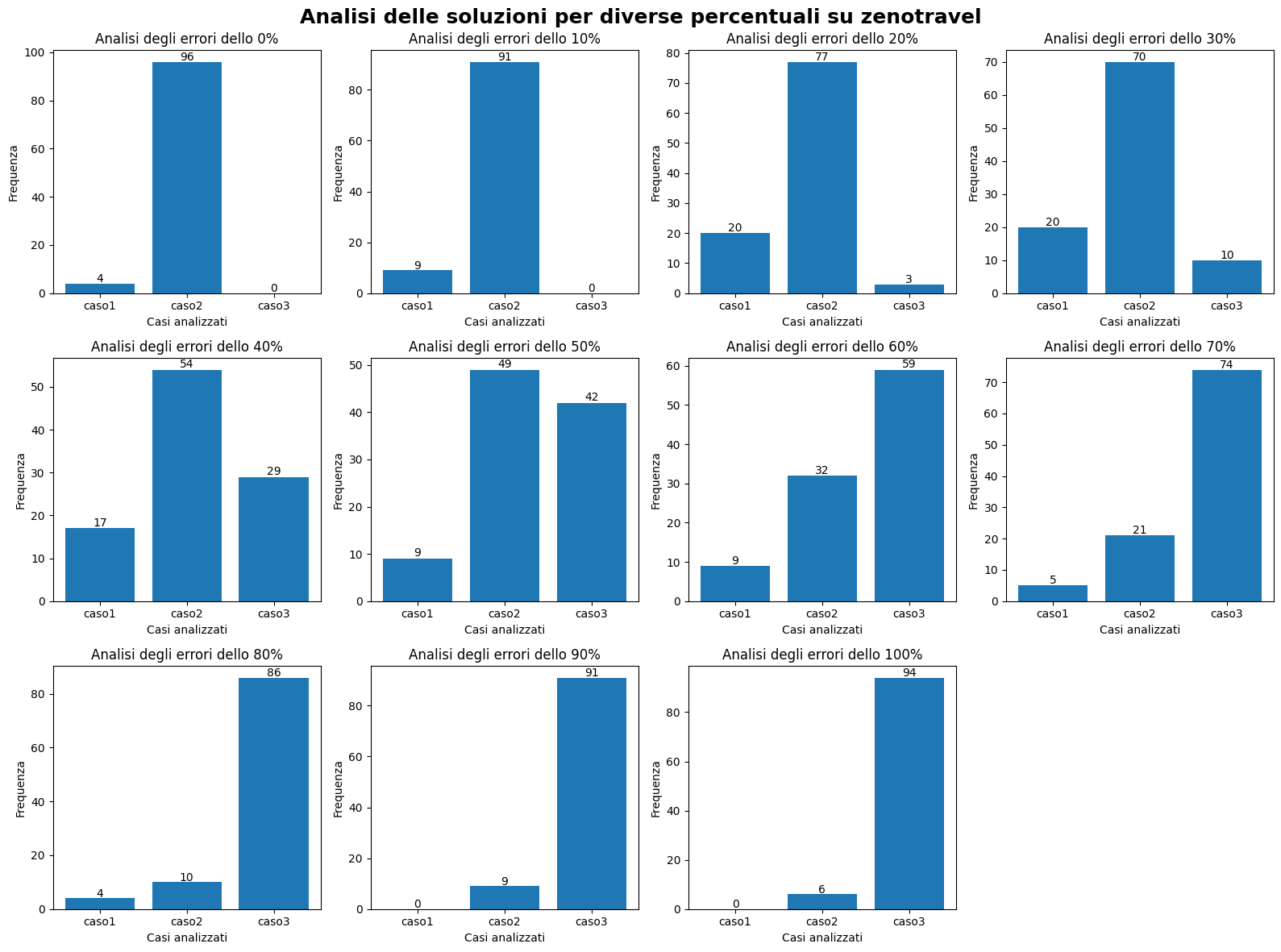
**Immagine che contiene testo, diagramma, schermata, Parallelo

Descrizione generata automaticamente** **Immagine che contiene testo, diagramma, schermata, Parallelo

Descrizione generata automaticamente**

**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

Descrizione generata automaticamente

Immagine che contiene testo, diagramma, schermata, Parallelo

Descrizione generata automaticamente Immagine che contiene testo, diagramma, schermata, Parallelo

Descrizione generata automaticamente

**metrica: calcola il numero di goal soddisfatti ma penalizzando sulla lunghezza (no\_penality = false)**

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

Immagine che contiene testo, diagramma, schermata, linea

Descrizione generata automaticamente Immagine che contiene testo, diagramma, schermata, linea

Descrizione generata automaticamente

Immagine che contiene testo, diagramma, schermata, linea

Descrizione generata automaticamente Immagine che contiene testo, diagramma, schermata, linea

Descrizione generata automaticamente

**metrica: calcola il numero di goal soddisfatti ma penalizzando sulla lunghezza (no\_penality = true)**

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

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

Descrizione generata automaticamente** **Immagine che contiene testo, diagramma, schermata, linea

Descrizione generata automaticamente** Immagine che contiene testo, diagramma, schermata, linea

Descrizione generata automaticamente Immagine che contiene testo, diagramma, schermata, linea

Descrizione generata automaticamente

**La soluzionGR 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: [0.312, 0.471, 0.563, 0.603, 0.623]

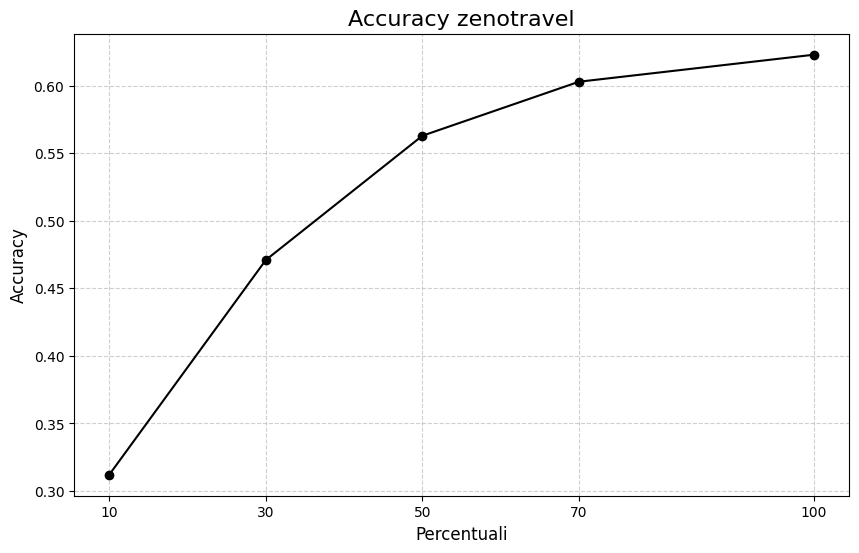


Immagine che contiene testo, diagramma, schermata, linea

Il contenuto generato dall'IA potrebbe non essere corretto.

**Driverlog**

Accuracy: [0.345, 0.48, 0.466, 0.483, 0.518]

Immagine che contiene testo, linea, Diagramma, diagramma

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.**

**Blocksworld**

Accuracy: [0.103, 0.244, 0.386, 0.501, 0.803]

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

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

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

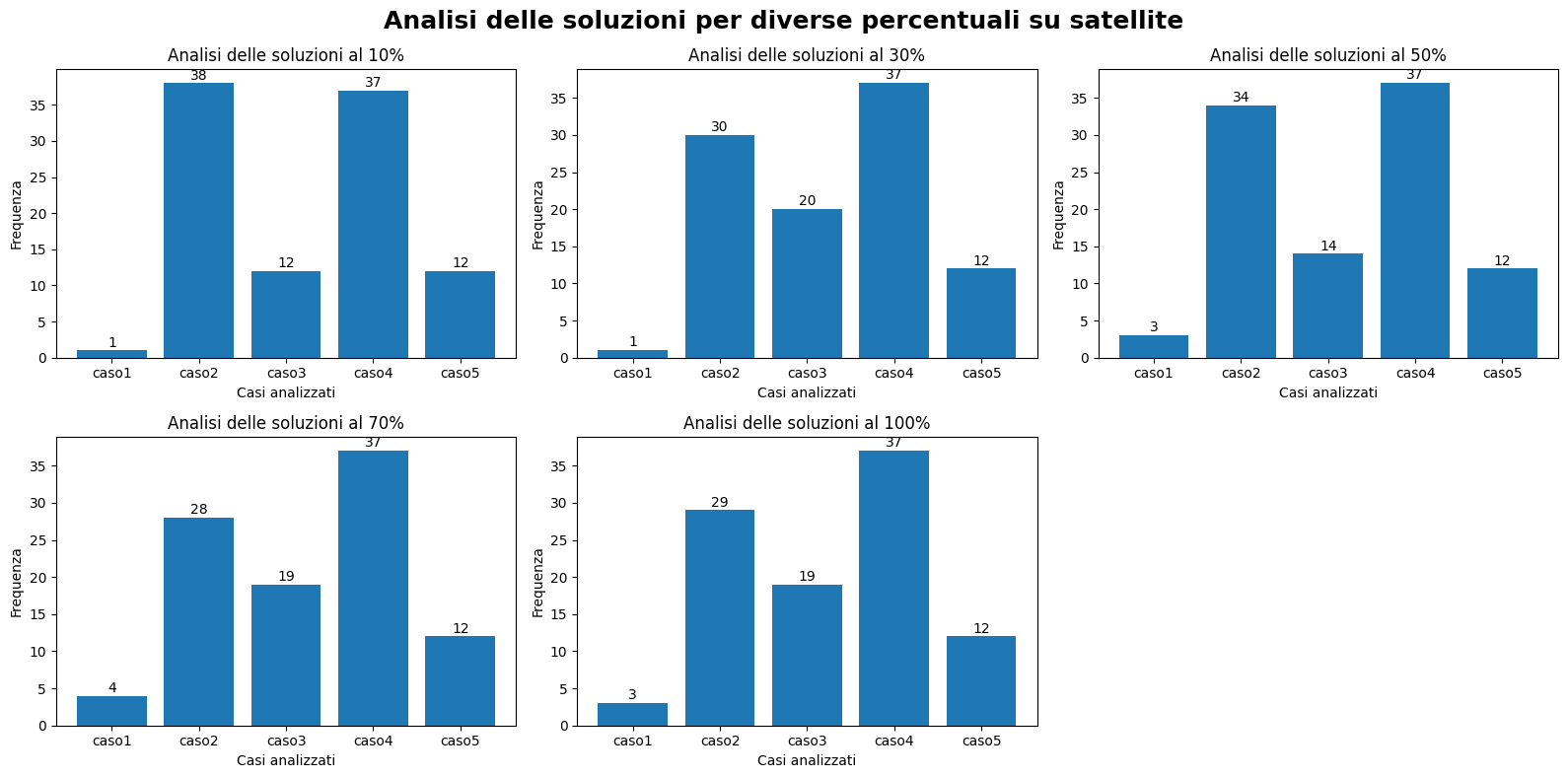
Il contenuto generato dall'IA potrebbe non essere corretto.**

**Satellite**

Accuracy: [0.219, 0.273, 0.228, 0.26, 0.269]

Immagine che contiene linea, Diagramma, diagramma

Il contenuto generato dall'IA potrebbe non essere corretto.

****