Supplementary Material

Domains description

We provide a very brief description of the eight well-known planning domains used in our experiments:

- **BLOCKSWORLD**. This domain relates to a robot that has to stack or unstack blocks, picking them up one at a time, to obtain a desired configuration of an available set of blocks. Our domain has configurations of single or multiple towers of blocks as both initial states and desired goal fluents.
- **DEPOTS**. This domain concerns actions for loading and unloading packages into trucks through hoists and moving them between depots. The goal is to have stacked packages at certain depots.
- **DRIVERLOG**. This domain involves driving trucks and delivering packages between locations. The complexity of this domain arises because trucks require drivers who must walk between trucks to drive them. The goal is to deliver a set of packages to their final destinations.
- **FLOORTILE**. This domain consists of robots that paint squares on a grid. The robots cannot traverse a painted square.
- **LOGISTICS**. In this domain, there are aircraft that can fly between cities, trucks that can move between locations within a city, and packages that can be loaded into/unloaded from trucks and aircraft. The goal is to deliver a set of packages to their final locations.
- **SATELLITE**. In this domain, a number of satellites collect image data of specific objects in the universe using different instruments with different modes.
- **VISITALL**. This domain consists of a single robot that has to visit squares on a grid.
- **ZENOTRAVEL**. This is another variant of a transportation domain in which passengers are embarked and disembarked in an aircraft that can fly between cities consuming fuel. The goal concerns transporting (moving) all passengers (aircraft) to their required destinations.

Size of the planning instances in the training domains

In this section, we give details about the number of objects involved in the planning instances used in the training phase of PLANGPT. For each type of object in a domain, we report its name, the minimum and maximum number of objects (min and max), and the total number of objects ($objs$); $objs$ indicates the number of all possible objects of a type that can be defined in a planning instance solvable by PLANGPT. We used these $[min, max]$ ranges to generate problems by varying the random seed.

- BLOCKSWORLD. $\{block : \{min : 3, max : 20, objs : 20\}\}$
- DEPOTS, { depot: $\{min : 1, max : 6, objs : 6\}$, distributor: $\{min : 1, max : 6, objs : 6\}$, truck: $\{min : 1, max : 6, objs : 6\}$, pallet: $\{min : 2, max : 15, objs : 15\}$, crate: $\{min : 2, max : 15, objs : 15\}$, hoist: $\{min : 2, max : 15, objs : 15\}$, }
- DRIVERLOG, { roads: $\{min : 1, max : 20, objs : 20\}$ driver: $\{min : 2, max : 8, objs : 8\}$ package: $\{min : 2, max : 25, objs : 25\}$ truck: $\{min : 2, max : 6, objs : 6\}$ }
- FLOORTILE { robot: $\{min : 1, max : 7, objs : 7\}$ grid: $\{min : 2x2, max : 7x7\}$ }
- LOGISTICS, { airplane: $\{min : 1, max : 8, objs : 8\}$, airport: $\{min : 2, max : 10, objs : 10\}$, location: $\{min : 1, max : 20, objs : 20\}$, city: $\{min : 2, max : 10, objs : 10\}$, truck: $\{min : 2, max : 10, objs : 10\}$, package: $\{min : 1, max : 18, objs : 18\}$ }
- SATELLITE { satellite: $\{min : 1, max : 10, objs : 10\}$ direction: $\{min : 1, max : 45, objs : 45\}$ mode: $\{min : 1, max : 5, objs; 5\}$ instrument: $\{min : 1, max : 29, objs : 29\}$ }
- VISITALL { rows: $\{min : 2, max : 11, objs : 10\}$ columns: $\{min : 2, max : 11, objs : 10\}$ % to visit: $\{min : 0.5, max : 1, objs : 11\}$ }
- ZENOTRAVEL, { aircraft: $\{min : 1, max : 5, objs : 5\}$, person: $\{min : 2, max : 20, objs : 20\}$, city: $\{min : 2, max : 18, objs : 18\}$, flevel: $\{min : 7, max : 7, objs : 7\}$ }

IPCScore-Agile and IPCScore-Quality

In this section, we illustrate for problems in a domain the scores *IPCScore-Quality* (IPCQ) and *IPCScore-Agile* (IPCA) as defined in the last International Planning Competition (IPC 2023).

- **IPCScore-Quality**: The score of a problem is the ratio C^*/C where C is the cost of the plan discovered by the model and C^* is the cost of a reference plan (the cheapest plan obtained by all models for that problem). The score of an unsolved problem is 0. The score of a model is the sum of its scores for all problems.
- **IPCScore-Agile**: The score of a problem on a solved task is 1 if the task was solved within 1 second and 0 if the task was not solved within the resource limits. If the problem was solved in T seconds ($1 \leq T \leq 300$) then its score is $1 - log(T)/log(300)$. The score of a model is the sum of its scores for all problems.

Domain	IPCQ GPT	IPCQ GNN	IPCA GPT	IPCA GNN	Cov GPT	Cov GNN	MT(s) GPT	MT(s) GNN	ML GPT	ML GNN
					IPC					
Blocks	**34.8**	34.3	**33.9**	25.9	**100.0**	**100.0**	**1.0**	4.4	**28.1**	28.9
Log	**16.0**	7.1	**13.3**	8.1	**53.3**	50.0	**3.8**	17.7	**31.1**	135.2
Visitall	17.2	**19.2**	**16.0**	9.0	95.0	**100.0**	**2.8**	23.7	48.2	**43.3**
					`Tset`					
Blocks	**6597.1**	1611.0	**6292.4**	1247.4	**100.0**	26.2	**1.3**	5.2	**39.6**	46.2
Log	**5125.1**	772.1	**4752.2**	791.7	**77.3**	21.6	**0.9**	21.1	**10.2**	80.6
Visitall	**6046.4**	6002.0	**5754.5**	3176.4	**100.0**	96.0	**2.4**	17.7	44.7	**42.5**

Table 4: Comparison of PLANGPT and GNN (Ståhlberg, Bonet, and Geffner 2022b) in terms of *IPCScore-Quality* (IPCQ), *IPCScore-Agile* (IPCA), coverage (Cov), Mean Time in seconds (s) (MT) and Mean Length (ML) of generated plan, for IPC and `Tset`.

Comparison with Graph Neural Networks

In Table 4 we show the comparison among PLANGPT (using the sampling generation strategy) and the Graph Neural Networks proposed in (Ståhlberg, Bonet, and Geffner 2022b) on IPC and `Tset` on the domains BLOCKSWORLD, LOGISTICS and VISITALL. Even if some of these problems are provided in training to GNNs, we show this comparison on IPC.

In terms of coverage, for BLOCKSWORLD, both models solve all IPC problems. In LOGISTICS, we obtain a slightly better percentage of coverage w.r.t. GNN (53.3 versus 50). However, in a modified version of LOGISTICS, changing its logical complexity from $C3$ to $C2$ (adding a predicate indicating in which city the objects, trucks and planes are located), GNN obtains 76.7% of coverage, outperforming PLANGPT. We tried applying this strategy to our model without improvement than the standard LOGISTICS. In VISITALL, GNN solves all IPC problems, while PLANGPT obtains a remarkable coverage of 95% of the IPC problems.

In terms of time, measured by the *IPCScore-Agile* (column IPCA), PLANGPT is faster than GNNs, obtaining better *IPCScore-agile* values for both the IPC test sets and for `Tset`. In terms of quality, measured by the *IPCScore-Quality* (column IPCQ), PLANGPT produces plans with better quality, obtaining better *IPCScore-Quality* values for all domains except VISITALL. We have already analysed the results obtained on `Tset` in Section *Comparison with the State of the Art*. In addition, here we report the *Mean Time* (MT) of the generation and *Mean Length* (ML) of the problem solution plan of both systems, which confirms the *IPCScores* results.

Training parameters of PLANGPT

In this section we report the training hyperparameters of PLANGPT on our domains:

- *num_decoder_layers* = 12
- *num_training_parameters* varies from $\{min : 82.70M; max : 83.03M\}$ depending on the dimension of the vocabulary provided for the domain
- *training_batch_size* = 4
- *eval_batch_size* = 4
- *learning_rate* = 5e-05 with no *weight_decay*
- *lr_scheduler* = *linear*
- *gradient_accumulation_steps* = 1
- *num_training_epochs* = 30
- *early_stopping* with Coverage Early Stopping and patience of 5 epochs
- *vocab_size* varies from $\{min : 119; max : 570\}$ depending on the planning domain

Extended Results

Table 5 shows the extended results for IPC$^-$ and `Tset` of PLANGPT with Coverage Early Stopping for each generation strategy (greedy, multibeam, and sampling). In this table, we use the following metrics: coverage, *IPCScore-Agile*, the *Mean Time* (MT) of the generation and the *Mean Length* (ML) of the plan.

We also report the error analysis for ZENOTRAVEL and DEPOTS. In the incorrect instances of ZENOTRAVEL, PLANGPT fails to satisfy all the goal fluents and violates the preconditions of the *board* action which, similarly to the action *load-truck* in LOGISTICS, requires understanding of the relationship between the objects involved. In DEPOTS, PLANGPT mostly violates the *unload* and *drop* actions.

Domain	Tset Size	Tset IPCA	Tset Coverage	Tset MT(s)	Tset ML	IPC Size	IPC IPCA	IPC Coverage	IPC MT(s)	IPC ML
					Greedy					
BLOCKSWORLD	6608	6480.2	99.5	0.9	39.0	35/35	34.8	100.0	0.7	29.3
DEPOTS	7041	5122.4	78.7	1.6	35.8	20/22	12.8	70.0	1.7	37.0
DRIVERLOG	7500	4214.9	68.4	3.1	86.0	20/20	15.3	80.0	1.2	29.9
FLOORTILE	6399	5238.5	94.4	2.3	61.0	20/20	15.9	100.0	3.3	81.4
LOGISTICS	6638	4308.4	66.1	0.8	20.1	30/30	11.4	40.0	1.3	27.7
SATELLITE	6505	4738.3	75.3	1.1	30.0	17/20	9.9	58.8	0.9	21.0
VISITALL	6565	5893.1	94.0	1.2	44.0	20/20	16.0	85.0	2.4	41.7
ZENOTRAVEL	7138	5407.1	82.7	1.7	39.8	18/20	16.0	94.4	1.4	29.6
					Sampling					
BLOCKSWORLD	6608	6292.5	100.0	1.2	38.2	35/35	34.2	100.0	0.9	28.1
DEPOTS	7041	5380.9	94.5	5.1	37.8	20/22	15.2	95.0	5.9	37.8
DRIVERLOG	7500	4585.7	96.5	9.6	85.7	20/20	16.0	95.0	4.5	46.5
FLOORTILE	6399	3939.3	99.6	13.2	62.7	20/20	10.4	100.0	16.9	82.8
LOGISTICS	6638	4752.2	77.3	1.6	21.7	30/30	13.0	53.3	3.6	35.2
SATELLITE	6505	5498.3	90.1	1.4	29.3	17/20	11.6	70.6	1,2	22.4
VISITALL	6565	5754.5	100.0	2.4	44.9	20/20	16.4	95.0	2.8	46.6
ZENOTRAVEL	7138	5733.3	94.7	2.7	39.5	18/20	16.0	100.0	2.2	29.8
					Multibeam					
BLOCKSWORLD	6608	6372.9	99.6	1.1	38.2	35/35	34.3	100.0	0.9	28.7
DEPOTS	7041	5282.3	85.4	2.2	37.3	20/22	11.9	70.0	3.5	49.6
DRIVERLOG	7500	4185.0	80.8	7.2	90.6	20/20	15.7	90.0	3.2	40.3
FLOORTILE	6399	4921.9	96.6	3.8	60.3	20/20	14.5	100.0	5.0	78.0
LOGISTICS	6638	4041.2	63.7	1.3	18.5	30/30	9.1	36.7	3.9	60.2
SATELLITE	6505	4729.9	78.3	1.5	29.1	17/20	8.7	47.0	1.3	19.9
VISITALL	6565	5543.5	97.8	2.9	53.5	20/20	16.0	95.0	3.3	53.3
ZENOTRAVEL	7138	5345.9	87.3	2.6	44.9	18/20	13.3	94.4	2.2	32.1

Table 5: Number of problems (Size), *IPCScore-agile* (IPCA), coverage, Mean Time(s) (MT) and Mean Length (ML) of generated plan, for Tset and IPC^{-} of PLANGPT in all the domains for greedy, multibeam and sampling generation.

Early stopping plots

In this section, for each domain, we show the plots of the cross-entropy loss function and the Coverage Early Stopping obtained by PLANGPT on the validation set for each training epoch. The black mark indicates the end of the training using the cross-entropy loss as early stopping metric. The red marker indicates the end of the training using Coverage Early Stopping .

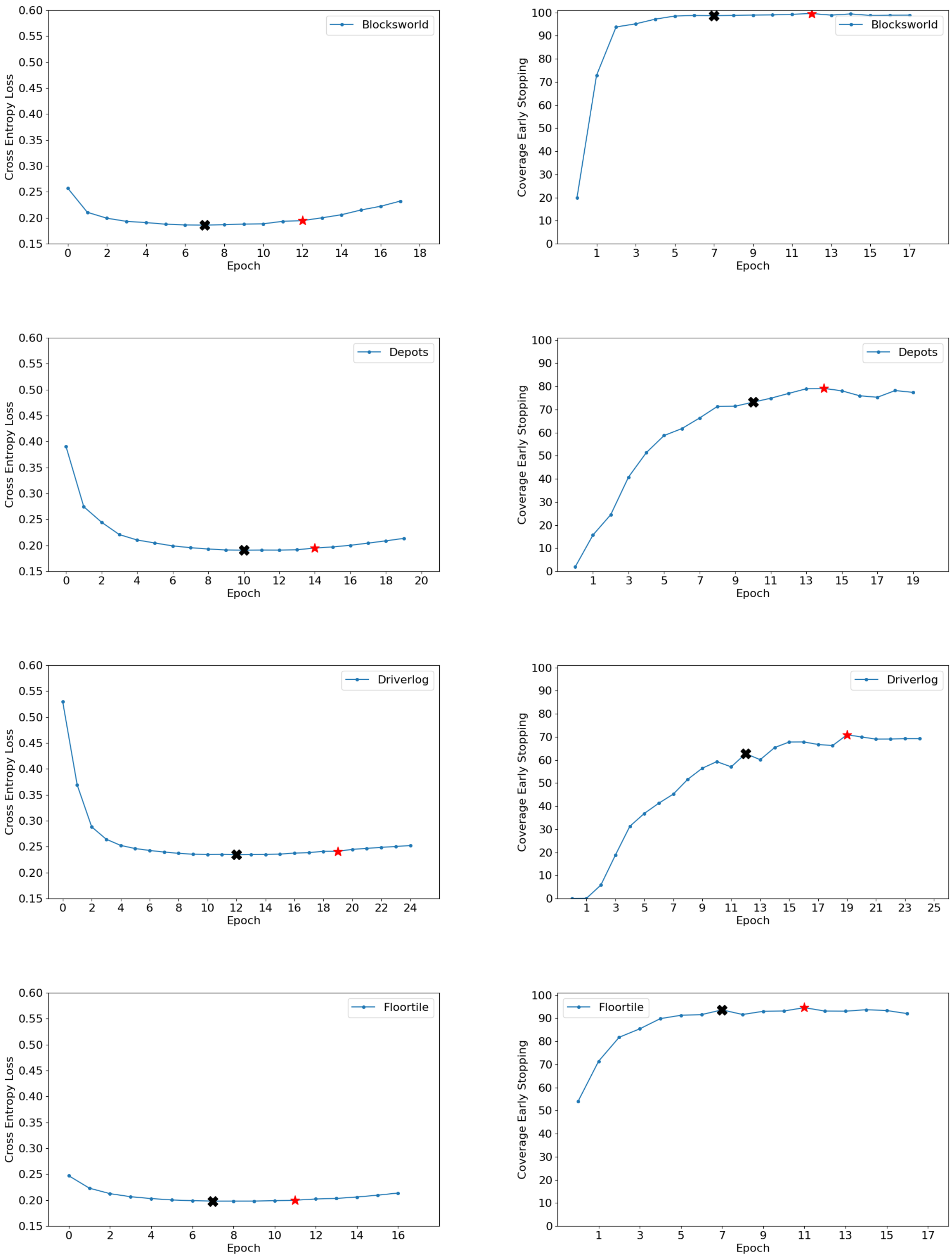

Figure 4: Cross Entropy Loss (on the left) and Coverage Early Stopping (on the right) for each epochs in the training phase of PLANGPT for BLOCKSWORLD, DEPOTS, DRIVERLOG and FLOORTILE domains.

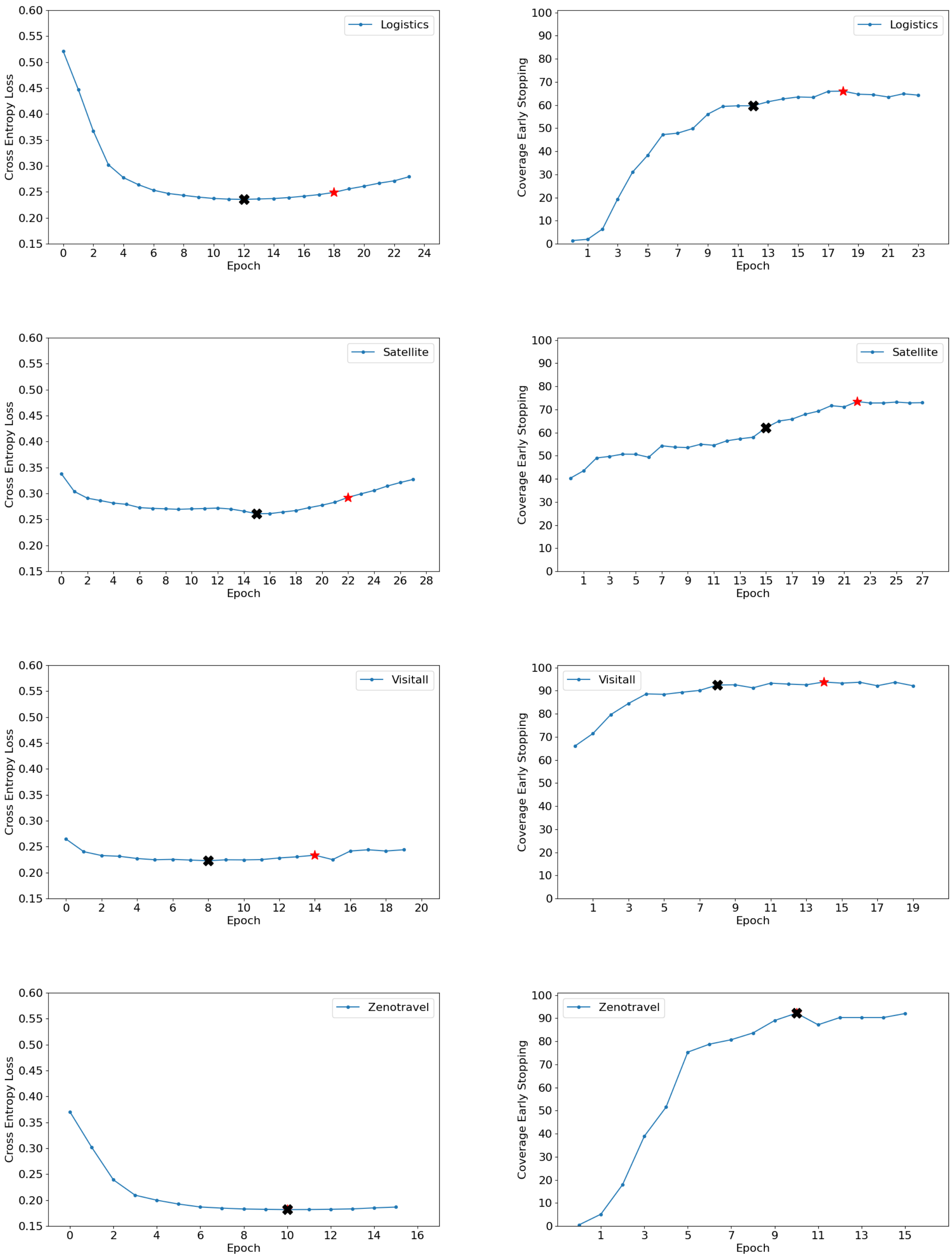

Figure 5: Cross Entropy Loss (on the left) and Coverage Early Stopping (on the right) for each epochs in the training phase of PlanGPT for LOGISTICS, SATELLITE, VISITALL and ZENOTRAVEL domains. ZENOTRAVEL stops on the same epoch with and without CES.