–Supplementary Material–
G2L: Semantically Aligned and Uniform Video Grounding via Geodesic and Game Theory

Hongxiang Li[1], Meng Cao[2,1], Xuxin Cheng[1], Yaowei Li[1], Zhihong Zhu[1], Yuexian Zou[1†]

[1]School of Electronic and Computer Engineering, Peking University

[2]International Digital Economy Academy (IDEA)

{lihongxiang, chengxx, zhihongzhu, ywl}@stu.pku.edu.cn; {mengcao, zouyx}@pku.edu.cn

A. Overview

In this supplementary material, we present the following.

- Axiomatic Properties of Shapley Value (Section B).
- Proof of Equation 10 (Section C).

B. Axiomatic Properties of Shapley Value

In this section, we mainly introduce the axiomatic properties of Shapley value. Weber *et al*. [1] have proved that Shapley value is the unique metric that satisfies the following axioms: *Linearity*, *Symmetry*, *Dummy*, and *Efficiency*.

Linearity Axiom. If two independent games u and v can be linearly merged into one game $w(\mathcal{U}) = u(\mathcal{U}) + v(\mathcal{U})$, then the Shapley value of each player $i \in \mathcal{N}$ in the new game w is the sum of Shapley values of the player i in the game u and v, which can be formulated as:

$$\phi_w(i|\mathcal{N}) = \phi_u(i|\mathcal{N}) + \phi_v(i|\mathcal{N}) \tag{1}$$

Symmetry Axiom. Considering two players i and j in a game v, if they satisfy:

$$\forall \mathcal{U} \in \mathcal{N} \setminus \{i, j\}, v(\mathcal{U} \cup \{i\}) = v(\mathcal{U} \cup \{j\}) \tag{2}$$

then $\phi_v(i|\mathcal{N}) = \phi_v(j|\mathcal{N})$.

Dummy Axiom. The dummy player is defined as a player without interaction with other players. Formally, if a player i in a game v satisfies:

$$\forall \mathcal{U} \in \mathcal{N} \setminus \{i\}, v(\mathcal{U} \cup \{i\}) = v(\mathcal{U}) + v(\{i\}) \tag{3}$$

then this player is defined as the dummy player. In this way, the dummy player i has no interaction with other players, *i.e.* $v(\{i\}) = \phi_v(i|\mathcal{N})$.

Efficiency Axiom. The efficiency axiom ensures that the overall reward can be assigned to all players, which can be formulated as follows:

$$\sum_{i \in \mathcal{N}} \phi_v(i) = v(\mathcal{N}) - v(\varnothing) \tag{4}$$

C. Proof of Equation 10

In this section, we provide detailed proof for Equation 10 in Section 3.5.2. The semantic Shapley interaction between moment x and query y in video V_i can be decomposed as follows:

$$\begin{aligned}
\mathfrak{I}([\mathcal{H}^i_{xy}]) &= \phi([\mathcal{H}^i_{xy}]|\mathcal{H}^i \setminus \mathcal{H}^i_{xy} \cup \{[\mathcal{H}^i_{xy}]\}) \\
&- \phi(\mathbf{h}^V_{ix}|\mathcal{H}^i \setminus \mathcal{H}^i_{xy} \cup \{\mathbf{h}^V_{ix}\}) \\
&- \phi(\mathbf{h}^Q_{iy}|\mathcal{H}^i \setminus \mathcal{H}^i_{xy} \cup \{\mathbf{h}^Q_{iy}\}) \qquad (5) \\
&= \mathop{\mathbb{E}}_{C}\{ \mathop{\mathbb{E}}_{\substack{\mathcal{U} \subseteq \mathcal{H}^i \setminus \mathcal{H}^i_{xy} \\ |\mathcal{U}|=C}} [f(\mathcal{U} \cup \mathcal{H}^i_{xy}) - f(\mathcal{U})]\} \\
&- \mathop{\mathbb{E}}_{C}\{ \mathop{\mathbb{E}}_{\substack{\mathcal{U} \subseteq \mathcal{H}^i \setminus \mathcal{H}^i_{xy} \\ |\mathcal{U}|=C}} [f(\mathcal{U} \cup \{\mathbf{h}_i x^V\}) - f(\mathcal{U})]\} \\
&- \mathop{\mathbb{E}}_{C}\{ \mathop{\mathbb{E}}_{\substack{\mathcal{U} \subseteq \mathcal{H}^i \setminus \mathcal{H}^i_{xy} \\ |\mathcal{U}|=C}} [f(\mathcal{U} \cup \{\mathbf{h}^Q_{iy}\}) - f(\mathcal{U})]\} \qquad (6) \\
&= \mathop{\mathbb{E}}_{C}\{ \mathop{\mathbb{E}}_{\substack{\mathcal{U} \subseteq \mathcal{H}^i \setminus \mathcal{H}^i_{xy} \\ |\mathcal{U}|=C}} [f(\mathcal{U} \cup \mathcal{H}^i_{xy}) - f(\mathcal{U} \cup \{\mathbf{h}^V_{ix}\}) \\
&- f(\mathcal{U} \cup \{\mathbf{h}^Q_{iy}\}) + f(\mathcal{U})\ ]\ \} \qquad (7)
\end{aligned}$$

References

[1] Robert J Weber. Probabilistic values for games. *The Shapley Value. Essays in Honor of Lloyd S. Shapley*, pages 101–119, 1988. 1