

# ソースコードの類似性

石 尾 隆<sup>†</sup> 山 本 哲 男<sup>††</sup>

ソースコード再利用やコードクローン検出などの研究には、プログラム中に含まれた「類似」したソースコード断片の判定・探索を行っているものがあるが、「類似」ソースコードの定義は研究ごとに異なっていることが多い。そこで、本ワークショップでは、研究者や開発者にとって合意の取れた共通の「類似」基準を明らかにすることを目指す。

ワークショップの共同作業では、まず、参加者の研究グループなどで開発しているプログラム解析ツールを使って、機械的に「類似」と判定されるようなソースコード断片群を集めておく。これらを参加者全員で、それぞれの個人の判断基準に従って「類似」しているかどうかを判定していく。その判定結果を集計し、参加者らが持つ判断基準について議論を行う。

得られた集計結果は参加者で共有し、今後、様々な研究で、ソースコードの類似性判定の正解集合として活用可能とする予定である。

---

<sup>†</sup> 大阪大学 大学院情報科学研究科  
Graduate School of Information Science and Technology, Osaka University

<sup>††</sup> 立命館大学 情報理工学部  
College of Information Science and Engineering, Ritsumeikan University