Block-based programming with Scratch community data: A position paper

Sayamindu Dasgupta
MIT Media Lab
Cambridge, MA 02142
Email: sayamindu@media.mit.edu

***Abstract*—In this position paper, I describe the rationale behind, and the early design of Scratch Data Blocks – a Scratch-based block-programming toolkit that allows Scratch users to program with public data from the Scratch online community.**

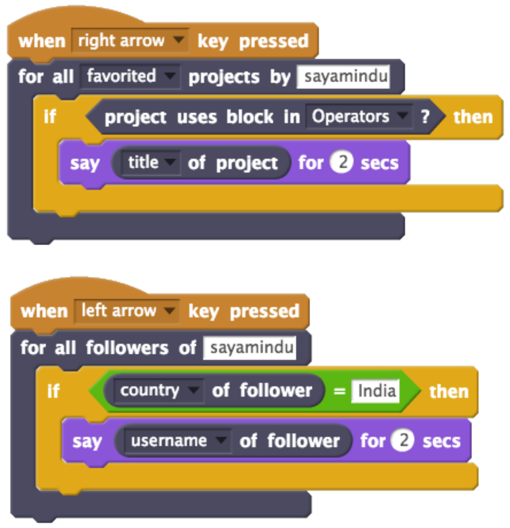
I. Introduction

A dominant trend in online communities is the collection and analysis of data from end-users. Though this data is then used for research, analytics, and marketing by the organization running the online-community, the users themselves rarely get to see this data. In many cases, users are not even aware that data is being collected about their activities in a given network – leading to what some scholars have termed as an "epistemological gap" [1] in privacy awareness. There is a similar trend in online learning communities and tools – though some recent developments show a positive change where designers of these communities and tools are actively engaging with the idea of presenting at least some data to community members and users through mechanisms such as dashboards, etc [2]–[4]. With the Scratch Data Blocks project, we aim to go one step further. Rather than having access to pre-programmed dashboards, with Scratch Data Blocks, young Scratch users will be able to create Scratch projects that can access, visualize, and analyze data about their own learning and participation (as well as their peers') in the Scratch community. Scratch Data Blocks will allow programmatic access to publicly available metadata about users and shared projects in the Scratch community. This would include not just social data (e.g. number of comments of a project, or the list of followers of a given user), but also, metadata about project code (e.g. number of blocks being used, or whether a certain type of block has been used or not).

With Scratch Data Blocks, we hope that Scratch users will find an engaging data-set (one that has been created by them and their peers), and utilize it to study and reflect upon not only their social participation, but also on their evolution as budding programmers.

II. Preliminary Design

In the initial design of Scratch Data Blocks, queries are expressed through "C-shaped" loop blocks. For example, projects can be retrieved through "foreach" queries of the form `for all shared projects by <<username>>` or `for all favorited projects by <<username>>` (Figure 1). Within these foreach loops or "C" blocks, context-sensitive reporter or predicate blocks can be used for retrieving metadata on the currently selected object (e.g. `title of project` or `country of user`). The available properties



```
when [right arrow] key pressed
for all [favorited] projects by [sayamindu]
    if <project uses block in [Operators] ?> then
        say (title of project) for (2) secs

when [left arrow] key pressed
for all followers of [sayamindu]
    if <(country of follower) = [India]> then
        say (username of follower) for (2) secs
```

Figure 1. Scratch Code using Scratch Data Blocks

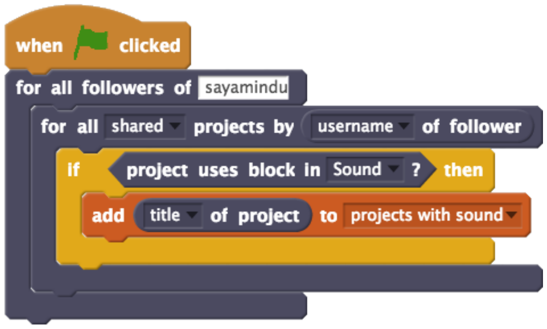
(e.g. country) for a given object-type (e.g. user) is listed in Table I. This approach, however, does represent a drawback of the current design, as Scratch has traditionally stayed away from context-sensitive blocks (with one notable exception of the `answer` reporter block to get user input). These context-sensitive blocks return a blank string, or false (in case of predicate blocks) when run outside of context (i.e. outside of the foreach loop).

Table I. Object-types and their properties

Object-type	Properties
User	username, bio, country
Project [a]	title, description, number of love-its, number of favorites, number of views

[a]The project object-type also has a predicate block for testing if a project uses a particular category of block, and another reporter block that returns the number of occurrences of a given block in a project's code.

In this design, though the loop or "C" blocks are tied to individual users (e.g. all projects shared by a *specific user*), it is possible to branch out in the Scratch community social graph through nested "C" blocks. For example, it is possible to enumerate projects shared by all the followers of a user by nesting a `for all shared projects by <<username>>` block inside of a `for all followers of <<username>>` block (Figure 2). This can be extended



```
when clicked
for all followers of sayamindu
  for all shared projects by username of follower
    if project uses block in Sound ? then
      add title of project to projects with sound
```

Figure 2. Nested loops to enumerate projects by all followers of a user

further to enumerate projects by friends-of-friends, though this process will quickly become slow on the client-side, and resource-consuming on the server-side, and there may be a need to impose a limit on the level of nesting that can be done for these kind of usage-scenarios.

It is worthwhile to note here that only public data is reported by the blocks, i.e. objects can be referred to or retrieved only if they are accessible publicly on the Scratch website. If a project gets unshared, or deleted, or if an user closes down his or her account, the relevant data will not be accessible via the blocks any more.

III. Future Directions

Over the next few months, I plan to start working with a selected group of Scratch community members to refine and iterate on the design described above. However, there are a few additions we hope to make to the toolkit before user-testing begins. A common phenomena within the Scratch online community are projects that celebrate community milestones (such as the sharing of the 10-millionth project, etc.). These projects are currently created manually, and there have been requests from the community in the past for access to community-wide statistics (e.g. number of projects currently shared, or the number of registered user). I plan to add reporter blocks for these global community statistics in the near-term future.

More global "queries" are also being considered, not for the initial launch, but for a subsequent release – for example, queries that enumerate projects currently in the homepage rows of the website, or users who have contributed to a given Scratch studio. A major open question here is that of performance – while it is possible in theory to have a block that allows the enumeration of all users in the community, or all projects in the community, it would entail addressing considerable architectural challenges on the server-side.

Longer term plans include exploring the possibility of write-access to the website – so that, for example, a Scratch project can generate an automatic "thank you" comment for anyone who clicks on "love-it" for the project. However, this sort of affordance, though extremely powerful, poses significant technical, as well as moderation challenges, and needs careful planning and thought before implementation.

IV. Acknowledgements

The author would like to thank Mitchel Resnick, Natalie Rusk, and other members of the Lifelong Kindergarten group at MIT, as well as Benjamin Mako Hill from the University of Washington for their suggestions and feedback during the conceptualization and prototype-design of this project. Financial support for this work came from the National Science Foundation (grant number 1417952).

References

[1] L. Floridi, "Big Data and Their Epistemological Challenge," *Philosophy & Technology*, vol. 25, no. 4, pp. 435–437, Dec. 2012. [Online]. Available: http://link.springer.com/article/10.1007/s13347-012-0093-4

[2] E. Duval, "Attention Please!: Learning Analytics for Visualization and Recommendation," in *Proceedings of the 1st International Conference on Learning Analytics and Knowledge*, ser. LAK '11. New York, NY, USA: ACM, 2011, pp. 9–17. [Online]. Available: http://doi.acm.org/10.1145/2090116.2090118

[3] J. L. Santos, S. Govaerts, K. Verbert, and E. Duval, "Goal-oriented Visualizations of Activity Tracking: A Case Study with Engineering Students," in *Proceedings of the 2Nd International Conference on Learning Analytics and Knowledge*, ser. LAK '12. New York, NY, USA: ACM, 2012, pp. 143–152. [Online]. Available: http://doi.acm.org/10.1145/2330601.2330639

[4] J. L. Santos, K. Verbert, S. Govaerts, and E. Duval, "Addressing Learner Issues with StepUp!: An Evaluation," in *Proceedings of the Third International Conference on Learning Analytics and Knowledge*, ser. LAK '13. New York, NY, USA: ACM, 2013, pp. 14–22. [Online]. Available: http://doi.acm.org/10.1145/2460296.2460301