

<https://github.com/neoneye/ARC-Interactive>

ARC-Interactive

How does humans solve ARC tasks?

Simon Strandgaard, 2024-mar-06

History

Interaction history - 266 steps



Press “space” for play

248 draw, x: 5 y: 20 color: 8
249 draw, x: 2 y: 15 color: 8
250 draw, x: 3 y: 15 color: 8
251 draw, x: 4 y: 15 color: 8
252 draw, x: 4 y: 13 color: 8
253 draw, x: 1 y: 12 color: 8
254 draw, x: 2 y: 12 color: 8
255 draw, x: 3 y: 12 color: 8
256 draw, x: 4 y: 12 color: 8
257 draw, x: 3 y: 13 color: 8
258 draw, x: 2 y: 11 color: 8
259 draw, x: 3 y: 11 color: 8
260 draw, x: 4 y: 11 color: 8
261 draw, x: 3 y: 10 color: 8
262 draw, x: 3 y: 3 color: 8
263 draw, x: 2 y: 3 color: 8
264 draw, x: 2 y: 2 color: 8
265 submit, correct
266 hide editor show overview

266 steps

Interaction history - 149 steps

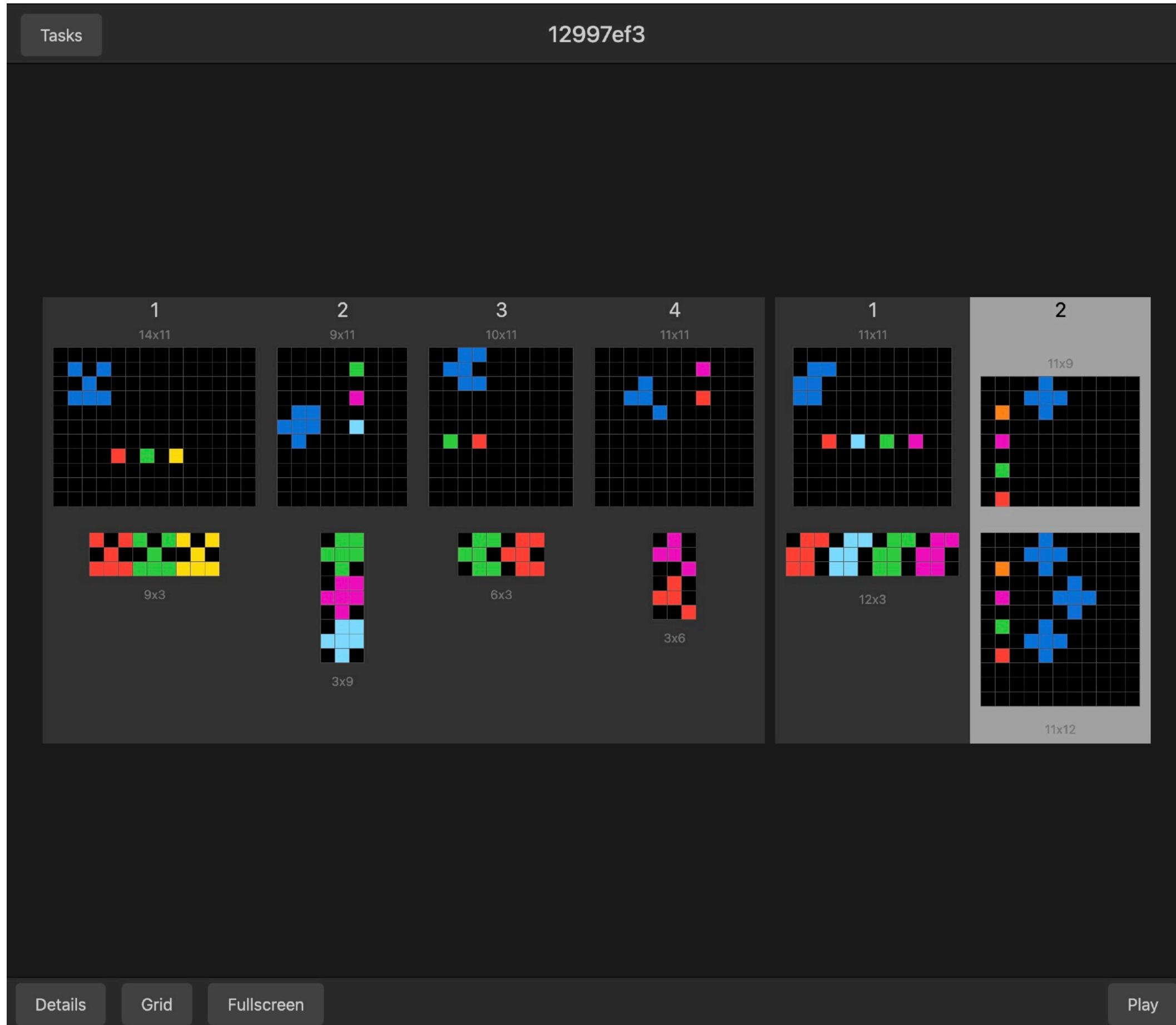


131	create selection update, x: 15 y: 3
132	create selection update, x: 15 y: 4
133	create selection update, x: 16 y: 4
134	create selection update, x: 16 y: 5
135	create selection update, x: 17 y: 5
136	create selection update, x: 17 y: 6
137	create selection update, x: 18 y: 6
138	create selection update, x: 18 y: 7
139	create selection update, x: 19 y: 7
140	create selection update, x: 19 y: 8
141	create selection update, x: 20 y: 8
142	create selection update, x: 20 y: 9
143	create selection update, x: 21 y: 9
144	create selection update, x: 21 y: 10
145	create selection update, x: 22 y: 10
146	show tool panel
147	crop, x: 12 y: 0 width: 11 height: 11
148	submit, correct
149	hide editor show overview

Press “space” for play

149 steps

Interaction history - 136 steps



Press “space” for play

118 flip x, x: 3 y: 0 width: 5 height: 3
119 create selection begin, x: 5 y: 0
120 create selection update, x: 5 y: 1
121 create selection update, x: 5 y: 2
122 create selection update, x: 6 y: 2
123 create selection update, x: 6 y: 3
124 create selection update, x: 6 y: 4
125 create selection update, x: 6 y: 5
126 create selection update, x: 6 y: 6
127 create selection update, x: 7 y: 6
128 create selection update, x: 7 y: 7
129 create selection update, x: 7 y: 8
130 create selection update, x: 7 y: 9
131 create selection update, x: 7 y: 10
132 create selection update, x: 7 y: 11
133 show tool panel
134 crop, x: 5 y: 0 width: 3 height: 12
135 submit, correct
136 hide editor show overview

136 steps

History json file

```
{  
  "fileType": "ARC-Interactive, history file, version 0.1.0",  
  "startTime": "2024-03-06T00:16:59Z",  
  "user": "anonymous",  
  "dataset": "1D-ARC",  
  "task": "1d_fill_6",  
  "summary": {  
    "history·count": 7,  
    "reveal·count": 0,  
    "start·over·count": 0,  
    "test·solved·count": 1,  
    "test·unsolved·count": 0,  
    "show·overview·count": 2  
  },  
  "history": [  ]  
}
```



Array with history events

Events in the history json file

“message”
is a human
readable
description.

```
{  
  "id": 3,  
  "wait": 966,  
  "message": "pick color 6",  
  "context": {  
    "action": "pick color",  
    "sameColor": false,  
    "imageHandle": "test-0-output",  
    "color": 6  
  },  
  "id": 4,  
  "wait": 882,  
  "message": "flood fill, x: 19 y: 0 color: 6",  
  "context": {  
    "action": "flood fill",  
    "sameImage": false,  
    "imageHandle": "test-0-output",  
    "x": 19,  
    "y": 0,  
    "color": 6,  
    "image": [ ]  
  },  
},
```

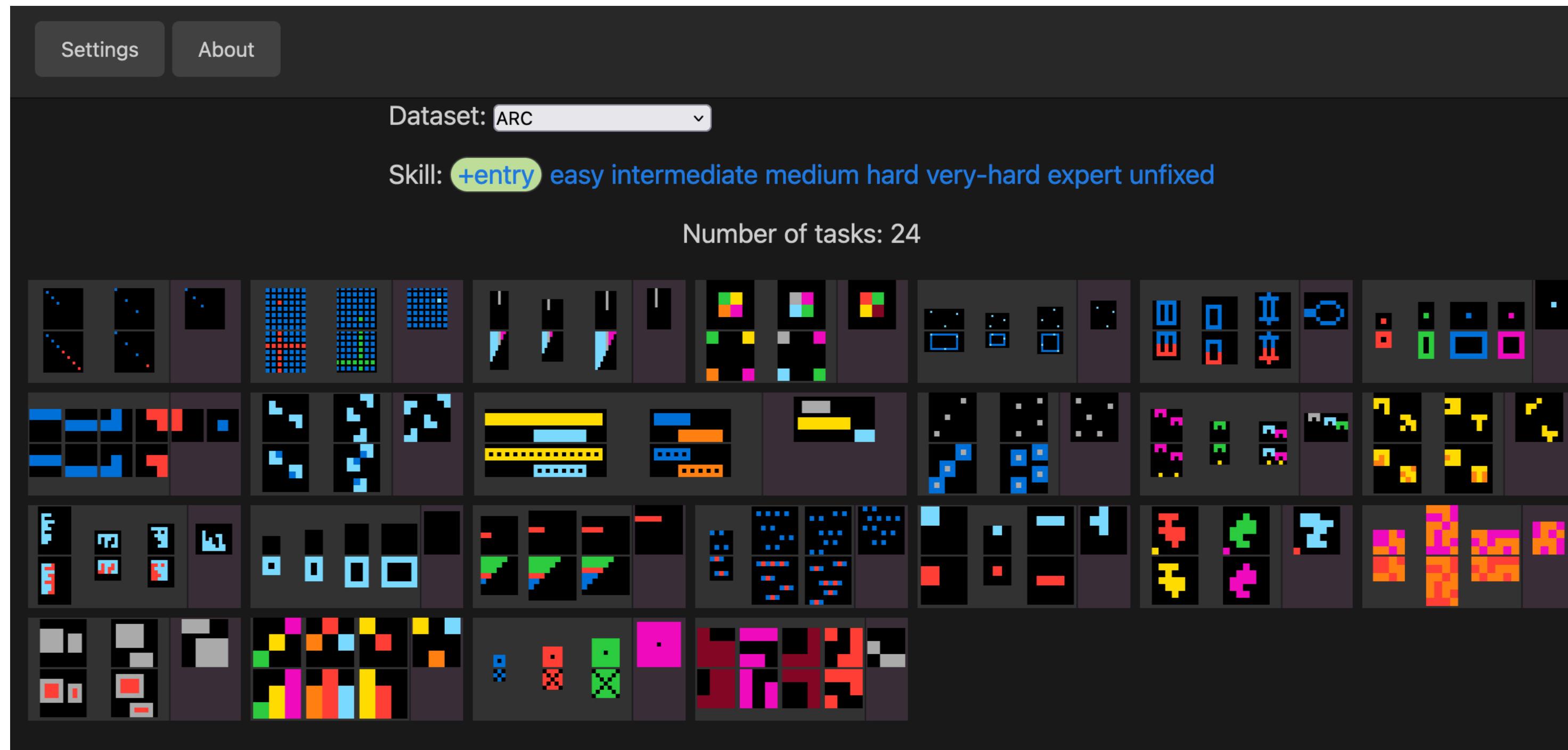
“context” provides more details than the human readable “message”.

Milliseconds since previous event.
May indicate if the human was thinking a lot or if it was easy.

Image snapshot after the event.

Gallery, Editor, Simple Mode

Tasks Gallery



Click on a thumbnail jumps to the editor in Simple Mode.
However in Advanced mode, it can jump other places.

Simple Mode

Thumbnails: 1 or 2 units wide.

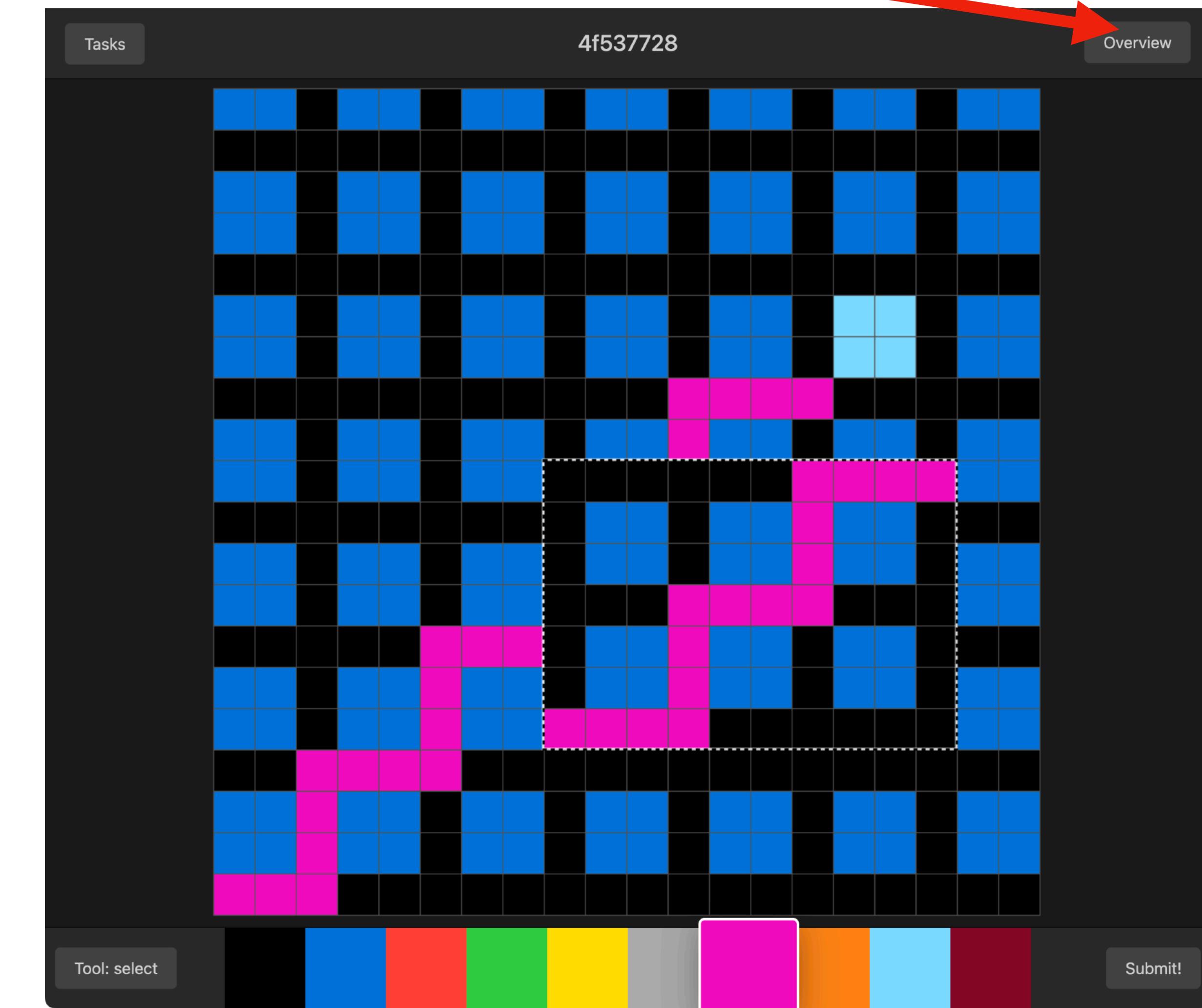
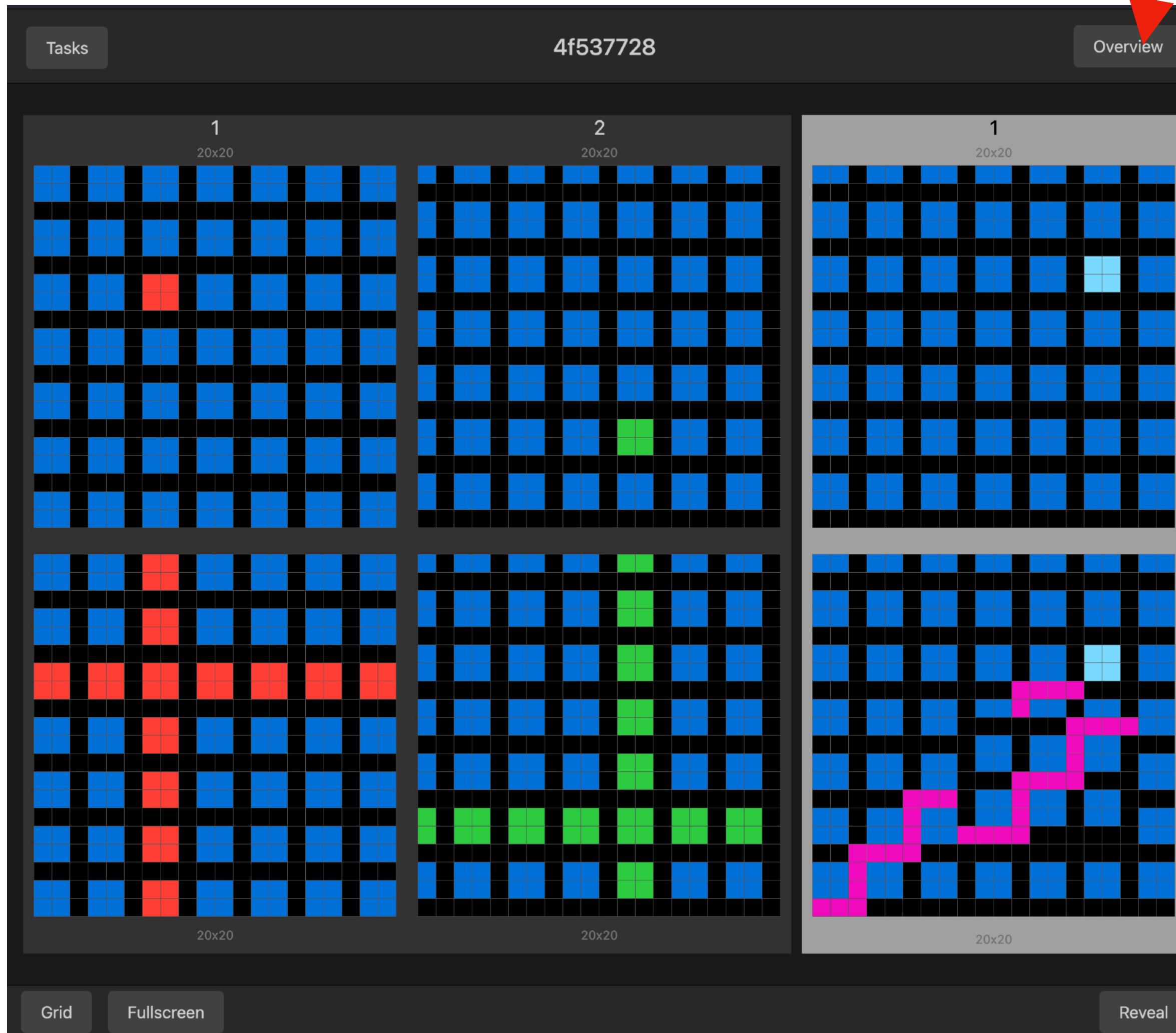
Dataset: ARC, MINI-ARC, ConceptARC

Skill: entry, easy. None of these requires resizing of the image.

Skill: intermediate...expert. These requires resizing the image.

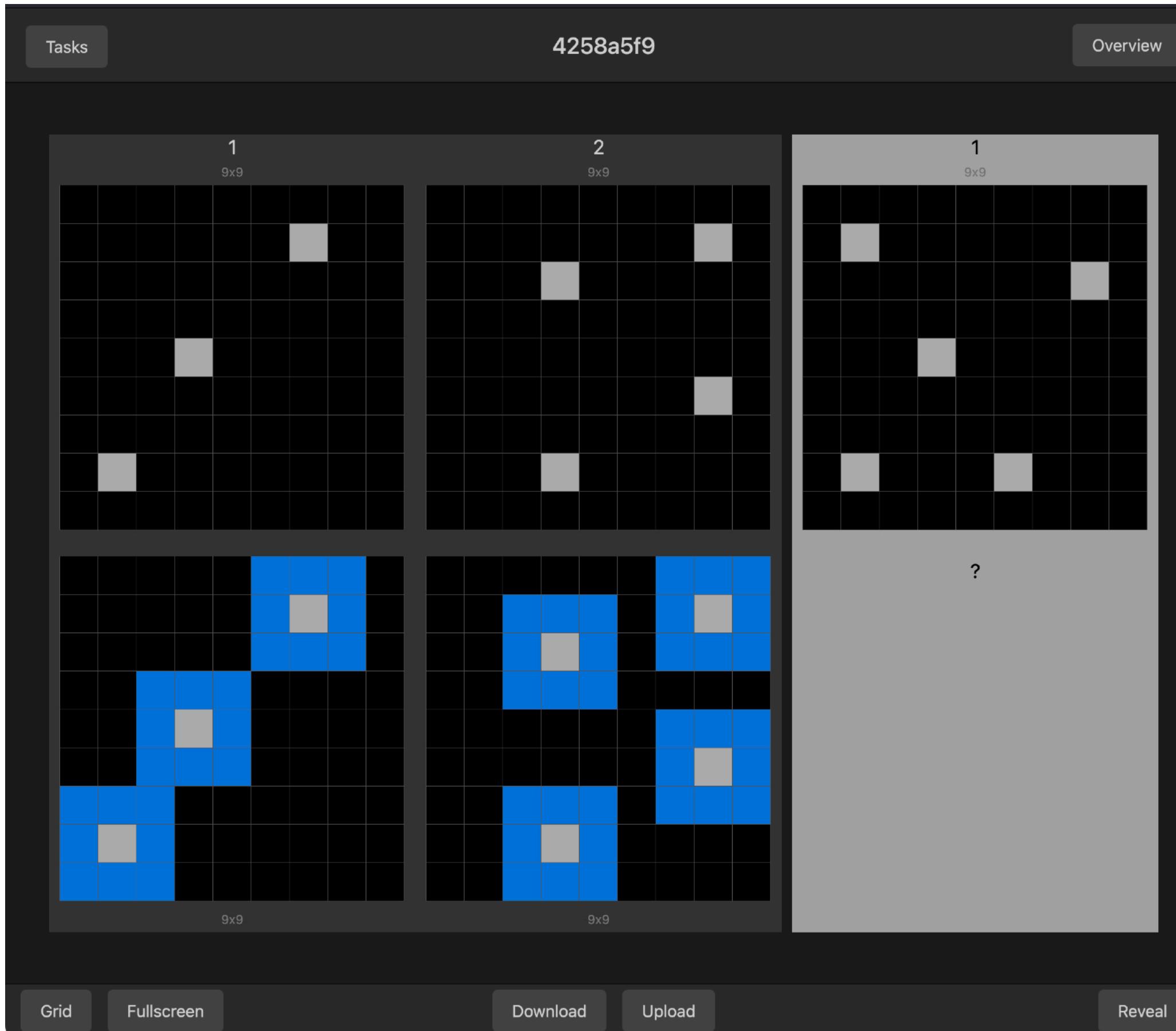
Click “Overview”, jumps between overview and edit

Overview

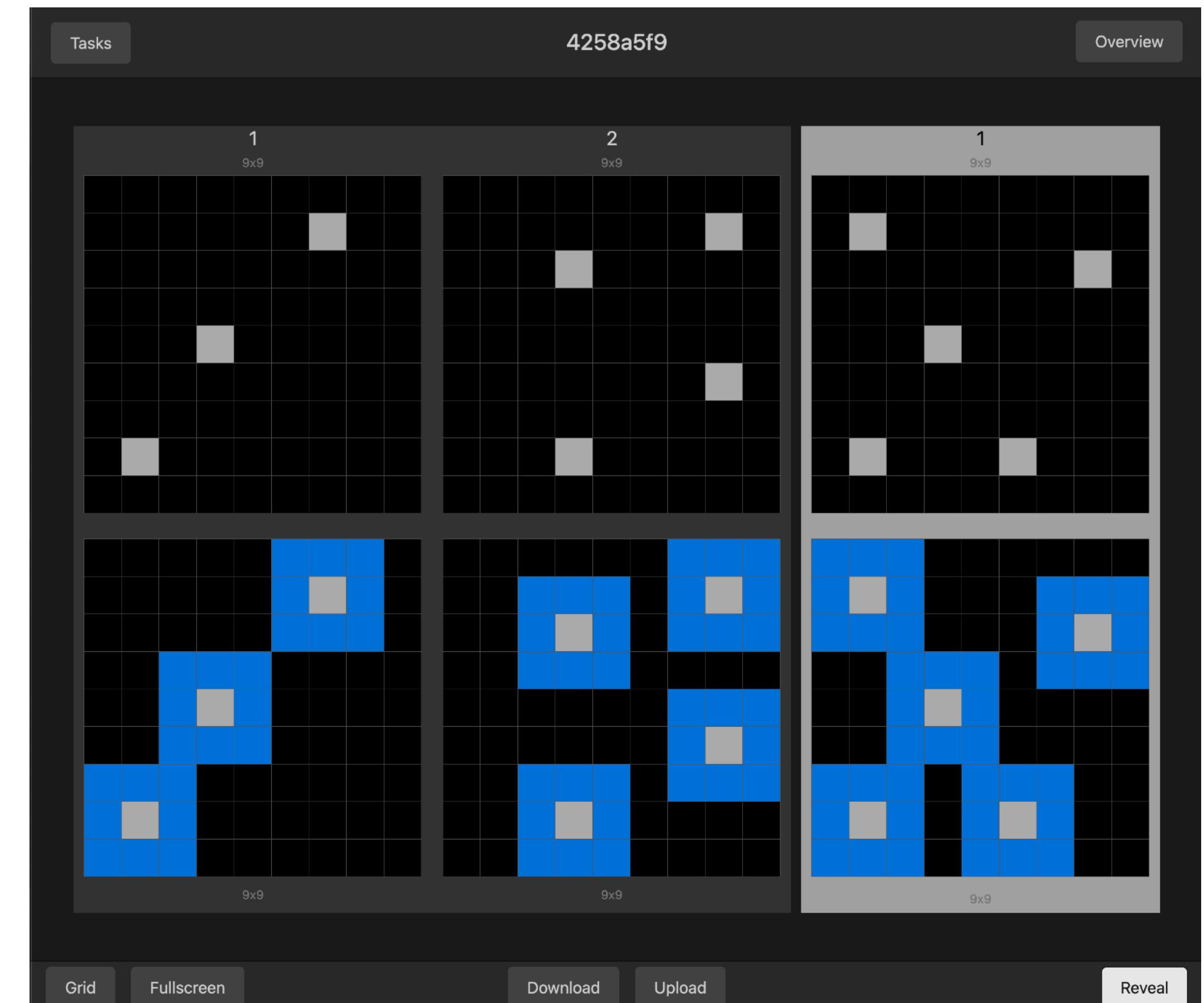


<https://neoneye.github.io/arc/edit.html?dataset=ARC&task=4f537728&filter=entry>

Hold down “Reveal” to see the expected output



Increments the counter
"reveal count":1
since it may disqualify the solution



Tools

Copy from input



Start over 😢可以更好

Play

See a replay
the history

Flip



Move

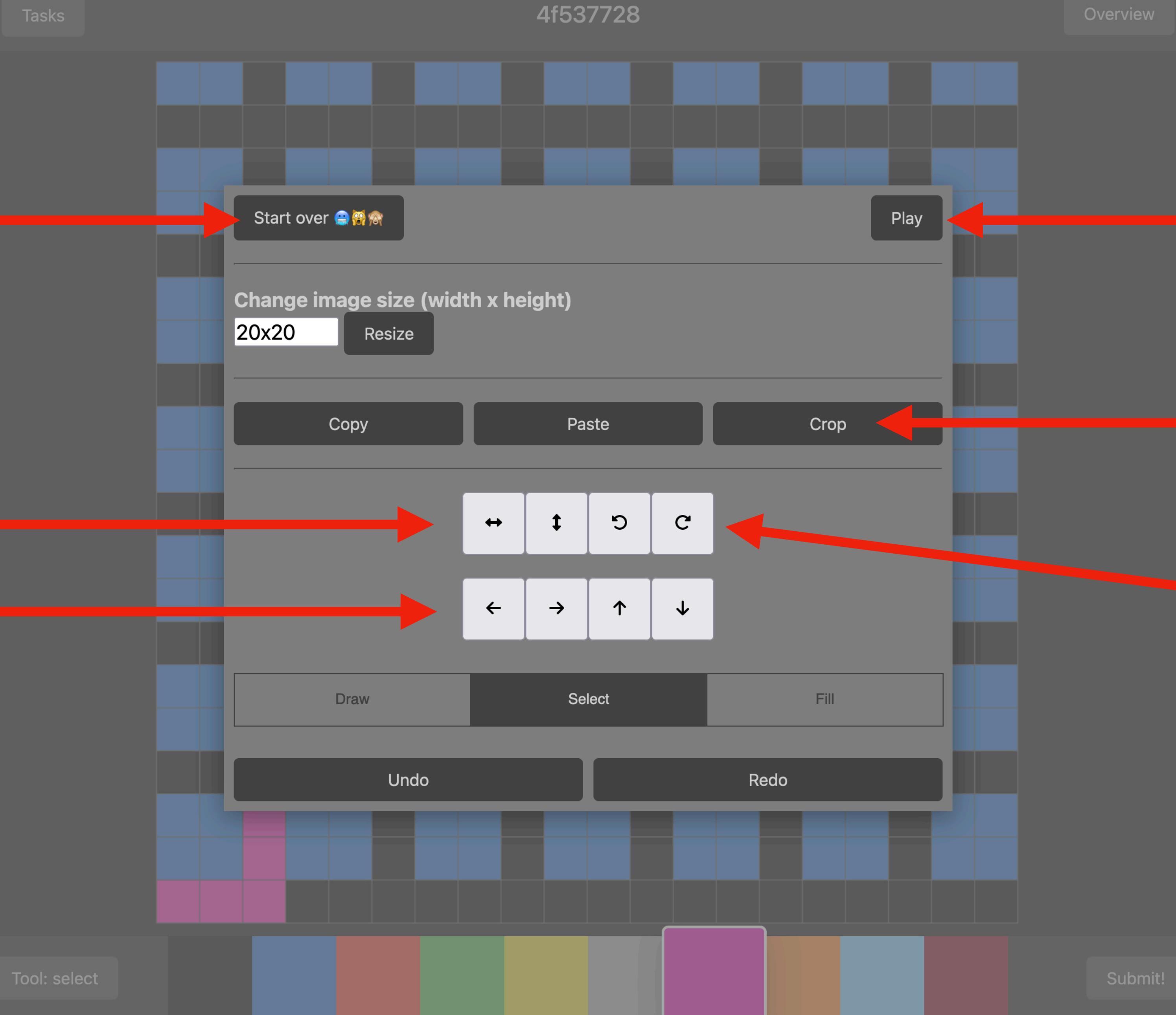


Crop out the
selected area

Crop

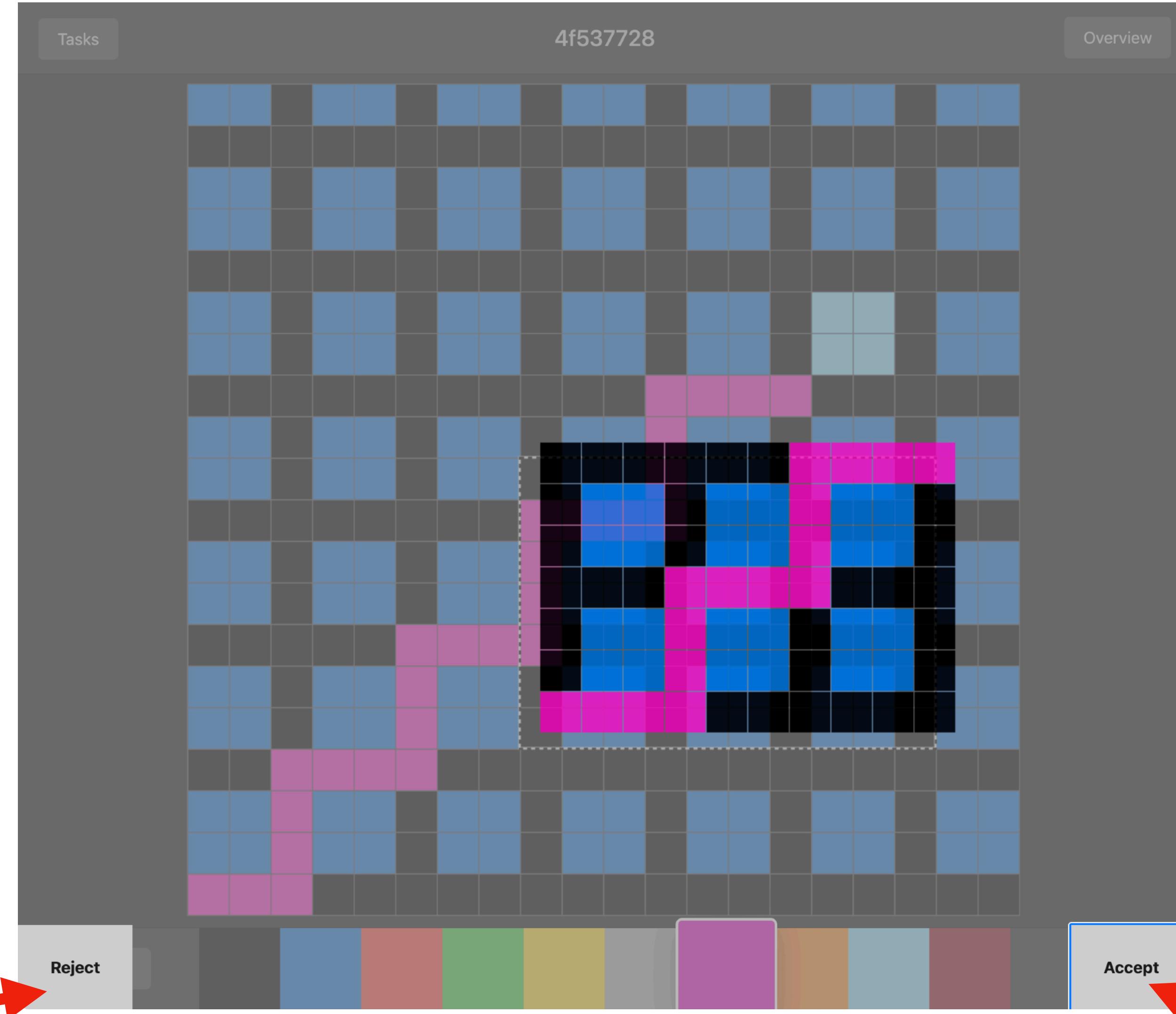


Rotate



Paste

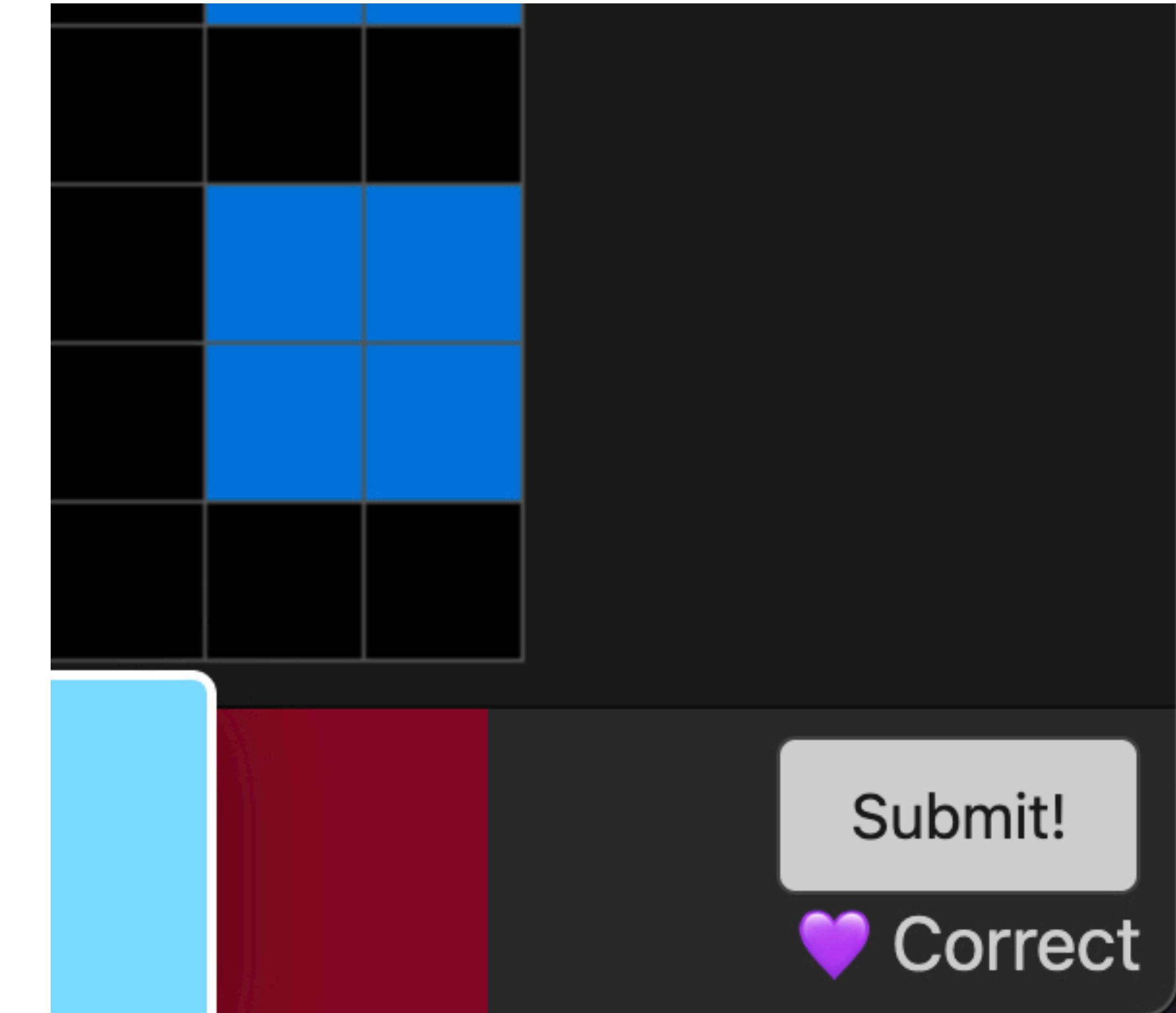
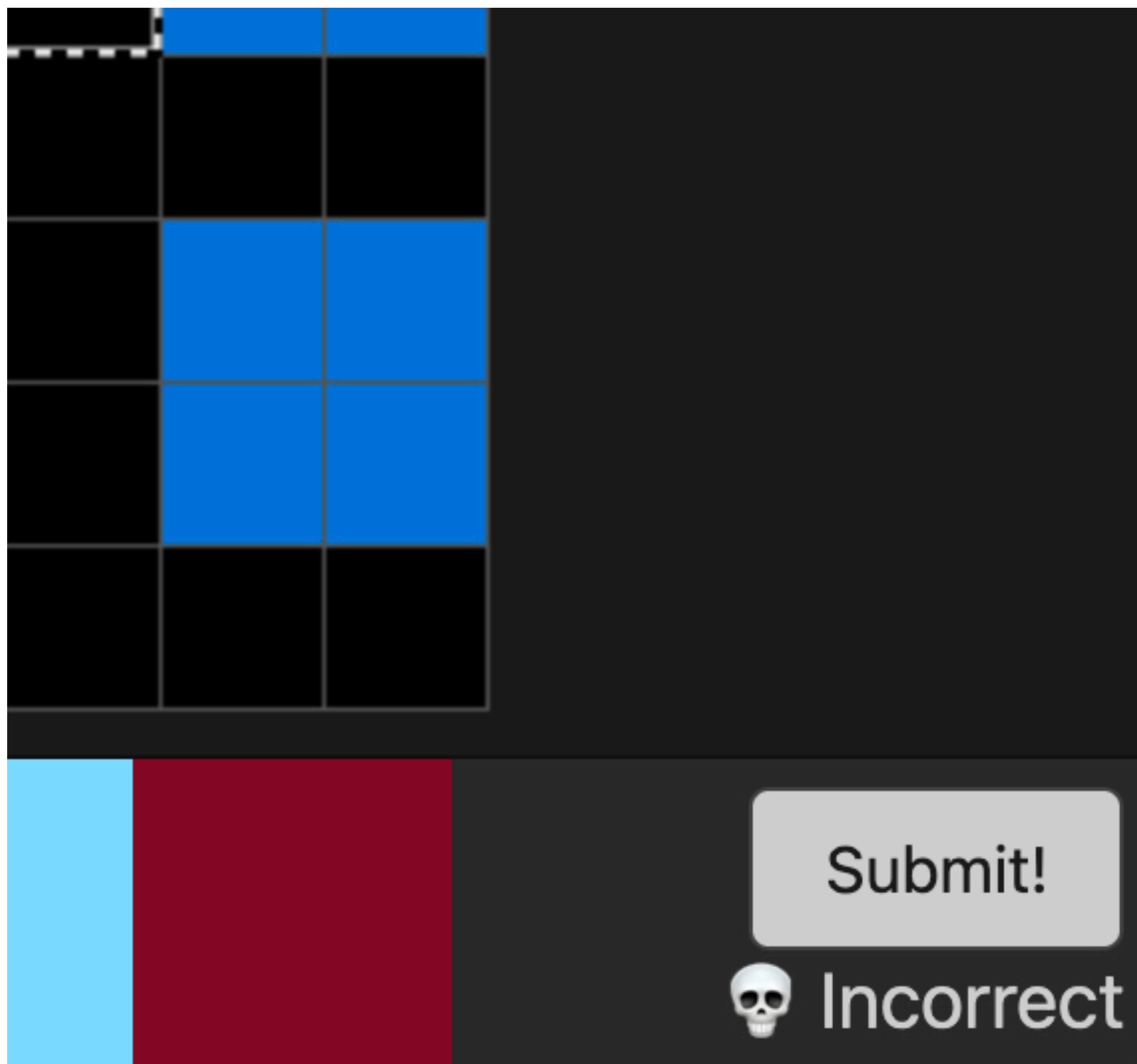
Move the semi-transparent paste layer around



Click “Reject” discards the paste

Click “Accept” inserts the paste

Click “Submit!” and Incorrect/Correct is shown



Increments these counters

"test solved count":1

"test unsolved count":0

Experimental Client Server project:
Submit sends history to server

Question for fchollet:
Regarding multiple test pairs.

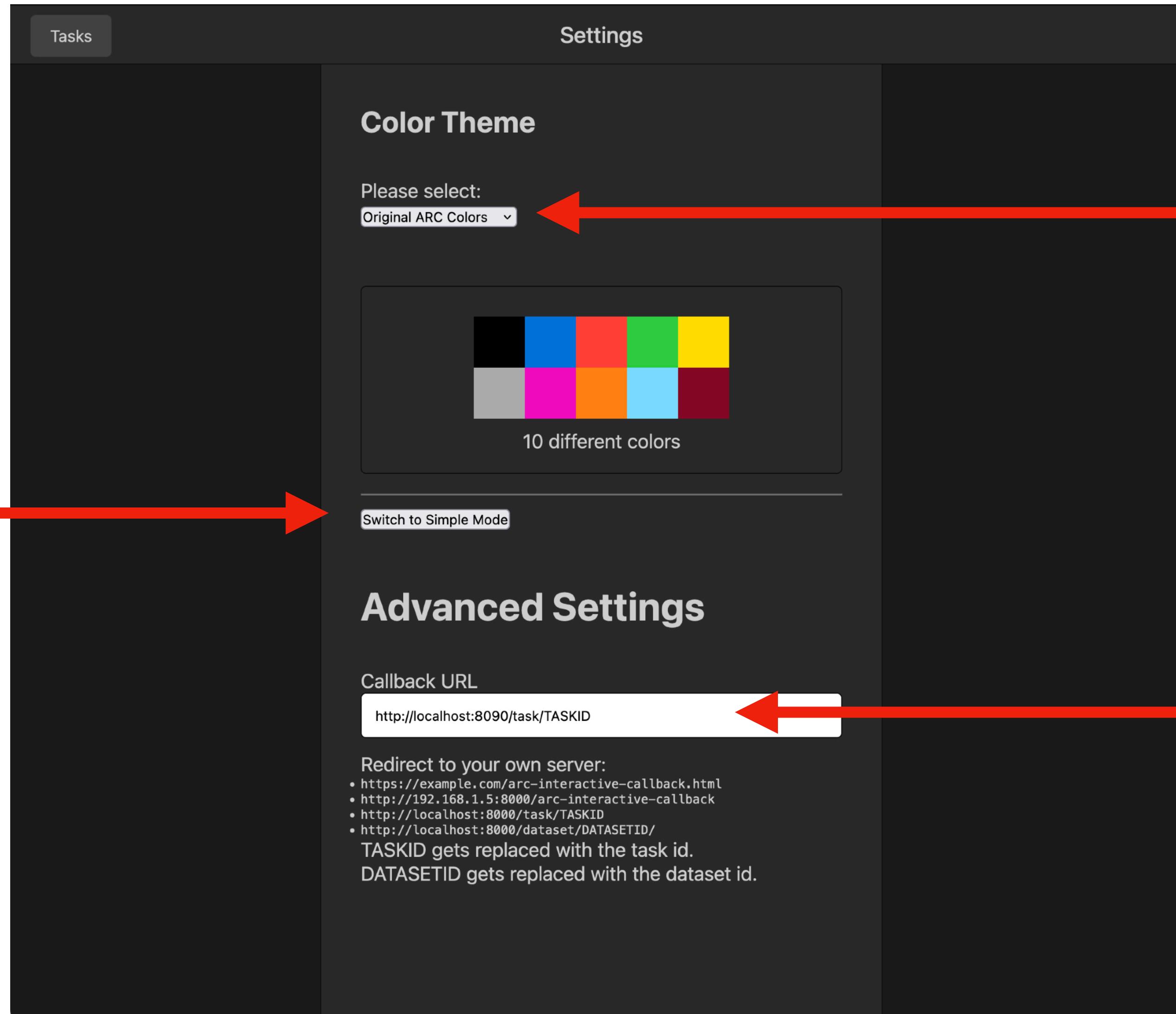
Scenario A: Currently the user can click “Submit!” and get feedback about that particular test pair.
Easy to determine which of the predictions are correct/incorrect.

Scenario B: Maybe the “Submit!” should only be clickable when the user have made predictions about all test pairs.
Hard to determine which of the predictions are correct/incorrect.

Gallery, Editor, Advanced Mode

Settings

Switch to
“Advanced mode”



Colorblindness

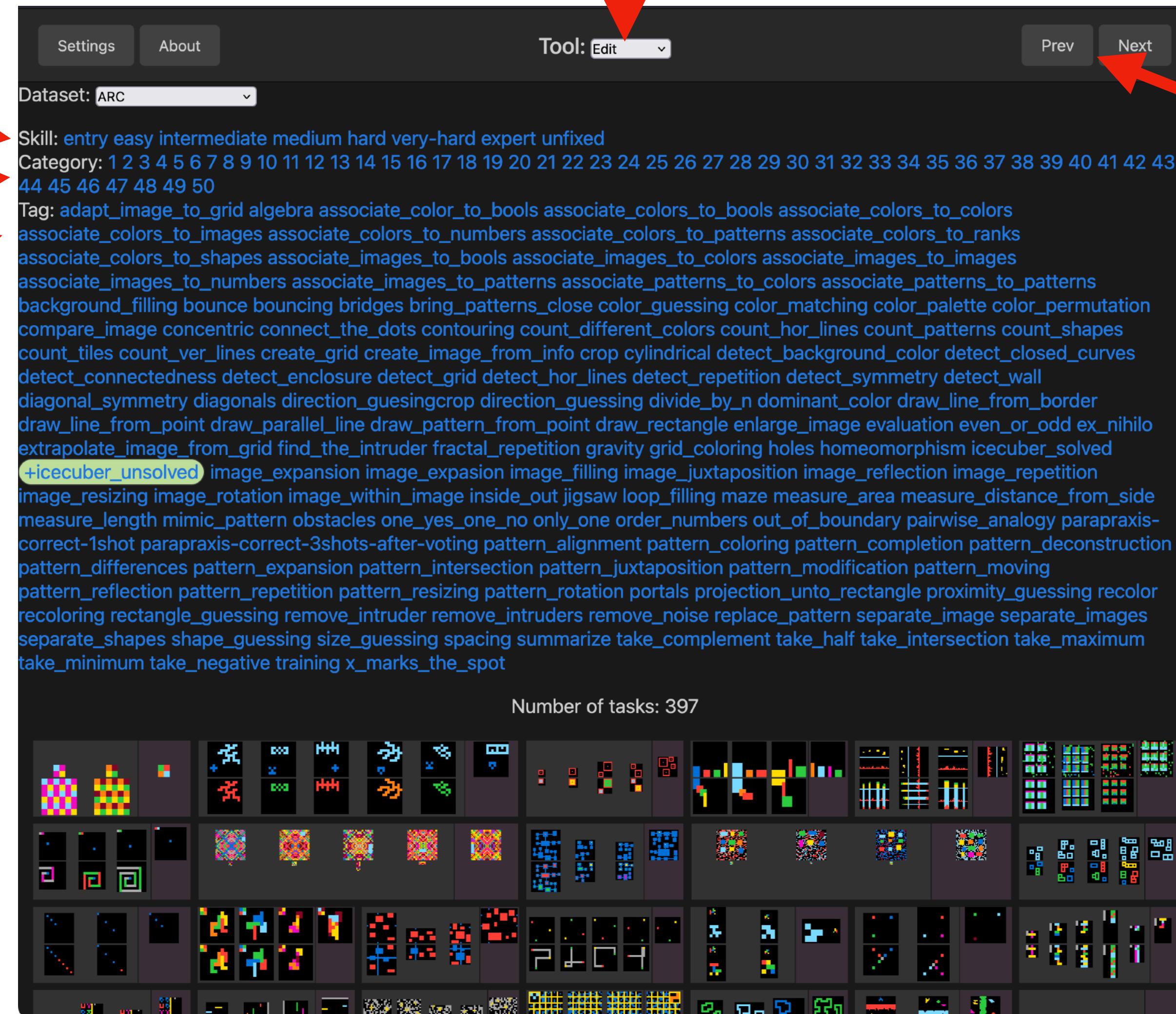
Callback

Tasks Gallery

Tool: Edit
what happens when
clicking a thumbnail

Advanced Mode

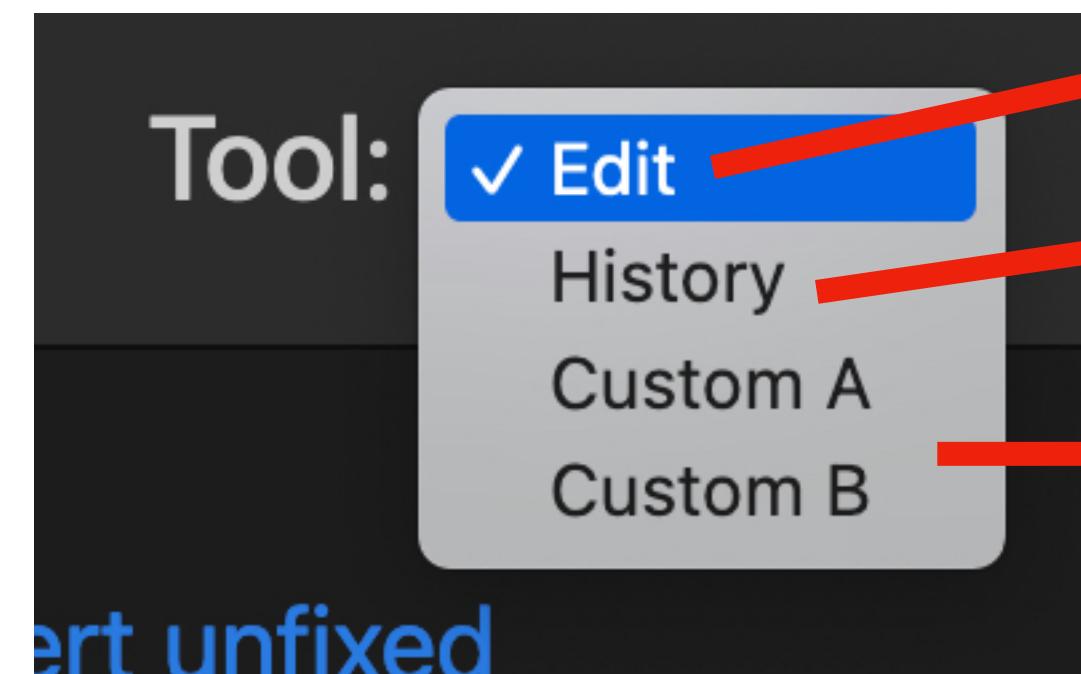
Skill
Category
Tag



Activate the
prev/next tag

Tasks Gallery

what happens when clicking a thumbnail



Jump to editor

Jump to history viewer

Jump to custom callback,
developers own web server
inspect class in detail
trigger the solver to run

Callback Configuration Required

You've reached a placeholder page because a callback URL has not been provided. Please configure the callback URL to proceed. [Settings](#).

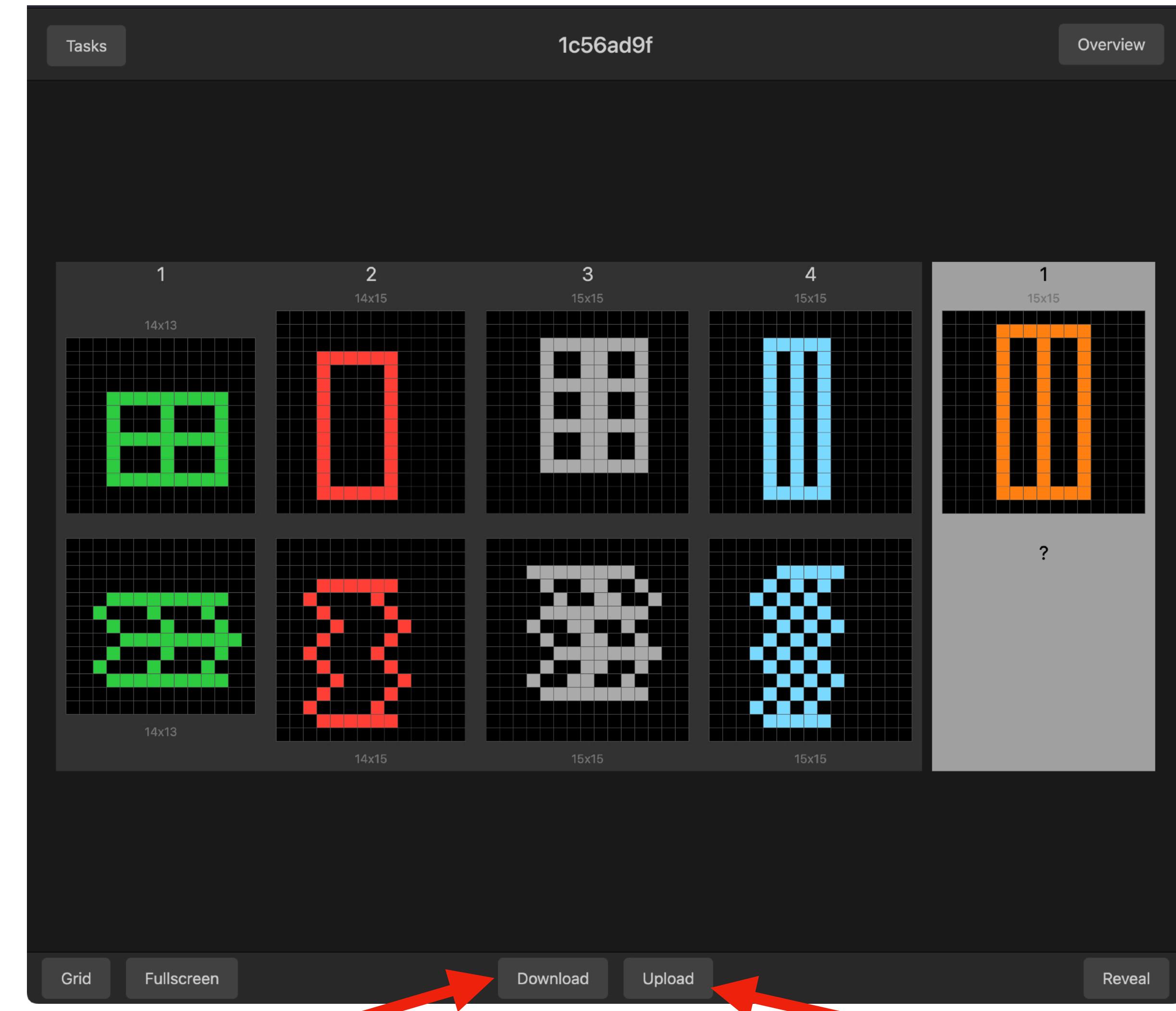
Parameters

action	custom-a
dataset	Mini-ARC
task	3_s_l6bksw4pewbthhupnrr

Download/Upload history file - kludgy

After the user have solved a task, click “**Download**” to get the corresponding history file.

It's hard to get friends to do this and send me the files. I have started coding on a server solution, that eliminates the “**Download**” button.



Download history json file

Upload history json file

To replay a previously solved task. Click “**Upload**” and select a history json file.

Client Server - work in progress

Review - step 1 of 3

127.0.0.1:5000/review

Tasks Review Delete Rejected Transfer Approved

ID	User ID	Timestamp	JSON size in bytes	Status	Link
8	1	2024-03-06 15:48:36.759403	98083	ReviewStatus.PendingReview	History
7	1	2024-03-06 15:48:32.046419	65437	ReviewStatus.PendingReview	History
6	1	2024-03-05 22:02:27.464145	24170	ReviewStatus.ReviewApproved	History
3	1	2024-03-05 18:12:58.866404	20278	ReviewStatus.ReviewRejected	History
2	1	2024-03-05 15:43:07.207102	16102	ReviewStatus.ReviewRejected	History

jumps to the history page

Approve Reject

0b17323b

2

15x15

1

15x15

15x15

15x15

15x15

15x15

Details Grid Fullscreen Play

Review - step 2 of 3

Play back how a user solved the task.
Determine if it's a keeper, or it should be discarded.

The image shows a composite view of a review interface and a history table. On the left, a 'Tasks' interface displays two 15x15 grids labeled '1' and '2'. Grid '1' contains blue squares, while grid '2' contains both blue and red squares. Above the grids are 'Approve' and 'Reject' buttons. On the right, a 'Review' interface shows a table of tasks. The columns are labeled 'Timestamp', 'JSON size in bytes', 'Status', and 'Link'. The table contains five rows of data:

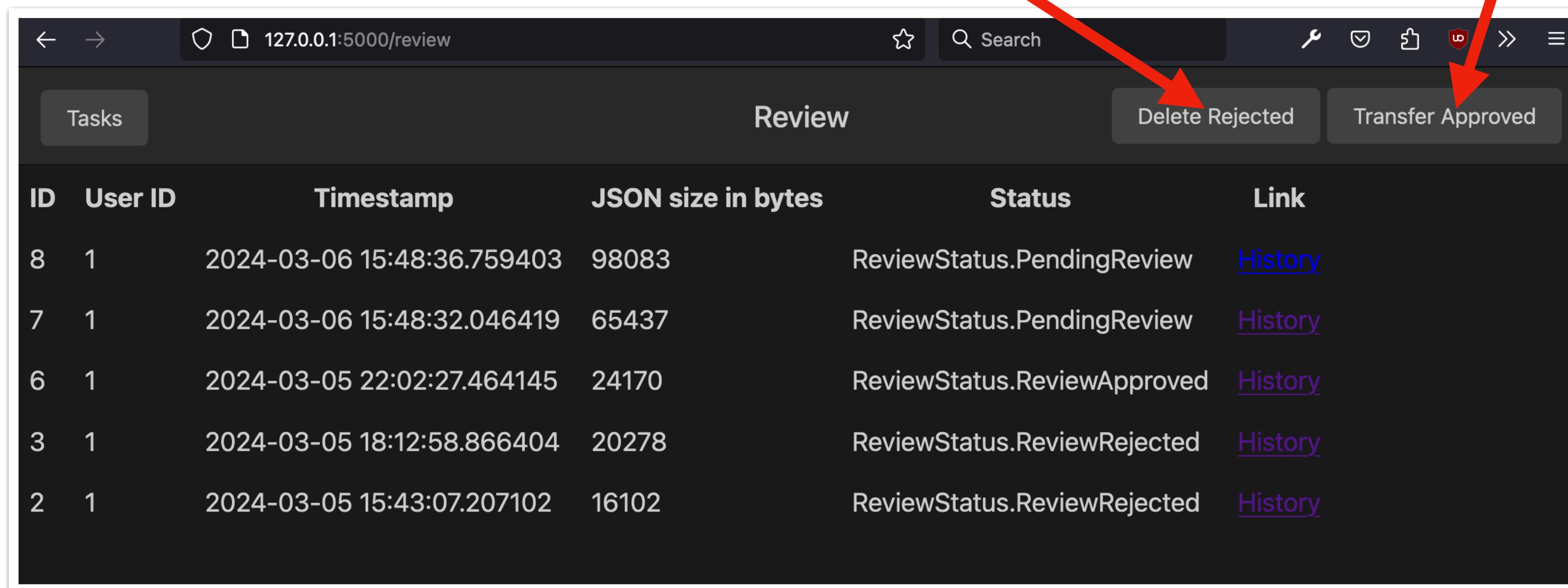
Timestamp	JSON size in bytes	Status	Link
48:36.759403	98083	ReviewStatus.PendingReview	History
48:32.046419	65437	ReviewStatus.PendingReview	History
02:27.464145	24170	ReviewStatus.ReviewApproved	History
2024-03-05 18:12:58.866404	20278	ReviewStatus.ReviewRejected	History
2024-03-05 15:43:07.207102	16102	ReviewStatus.ReviewRejected	History

A large red arrow points from the 'Approve' button in the top-left interface to the 'Transfer Approved' button in the bottom-right history table.

Review - step 3 of 3

click “Delete Rejected”
and the rejected history data
gets deleted.

click “Transfer Approve”
and the approved history data
are committed to the repository:
“ARC-Interactive-History-Dataset”



A screenshot of a web browser window titled "Review". The URL bar shows "127.0.0.1:5000/review". The main content area is a table with the following columns: ID, User ID, Timestamp, JSON size in bytes, Status, and Link. There are five rows of data:

ID	User ID	Timestamp	JSON size in bytes	Status	Link
8	1	2024-03-06 15:48:36.759403	98083	ReviewStatus.PendingReview	History
7	1	2024-03-06 15:48:32.046419	65437	ReviewStatus.PendingReview	History
6	1	2024-03-05 22:02:27.464145	24170	ReviewStatus.ReviewApproved	History
3	1	2024-03-05 18:12:58.866404	20278	ReviewStatus.ReviewRejected	History
2	1	2024-03-05 15:43:07.207102	16102	ReviewStatus.ReviewRejected	History

Q&A

Ideas for “Submit!”

Idea:

“Submit!” with a **confidence score**.

Unsure or very confident.

Is there a correlation between the history and the score.

Idea:

“Submit!” with a **comment textfield**.

Before: What kind of observations does the user make, before making any changes to the drawing.

After: What does the user say after having solved the task.

Ideas for the “Review” UI

Idea:

Assign **labels** to history data:
novel, advanced, beginner, made-a-mistake-but-fixed-it.

Idea:

Is the solution **above/below baseline**.
Does the user do it in fewer steps or more steps.