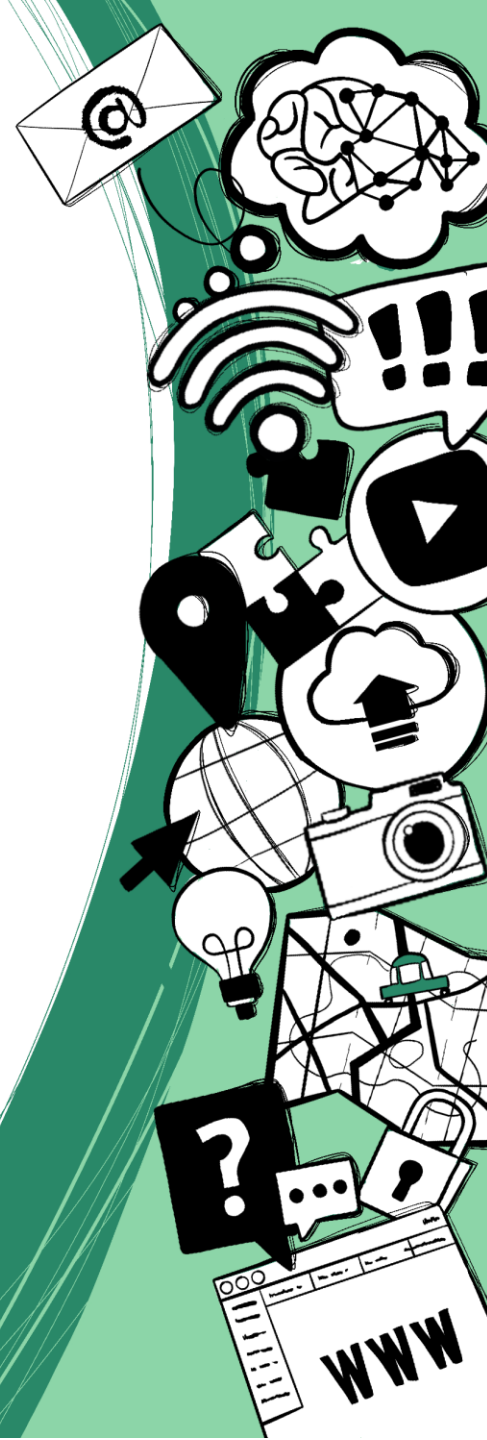
ENARIS

Labirintus



mobilis
interaktív kiállítási központ

TU Graz

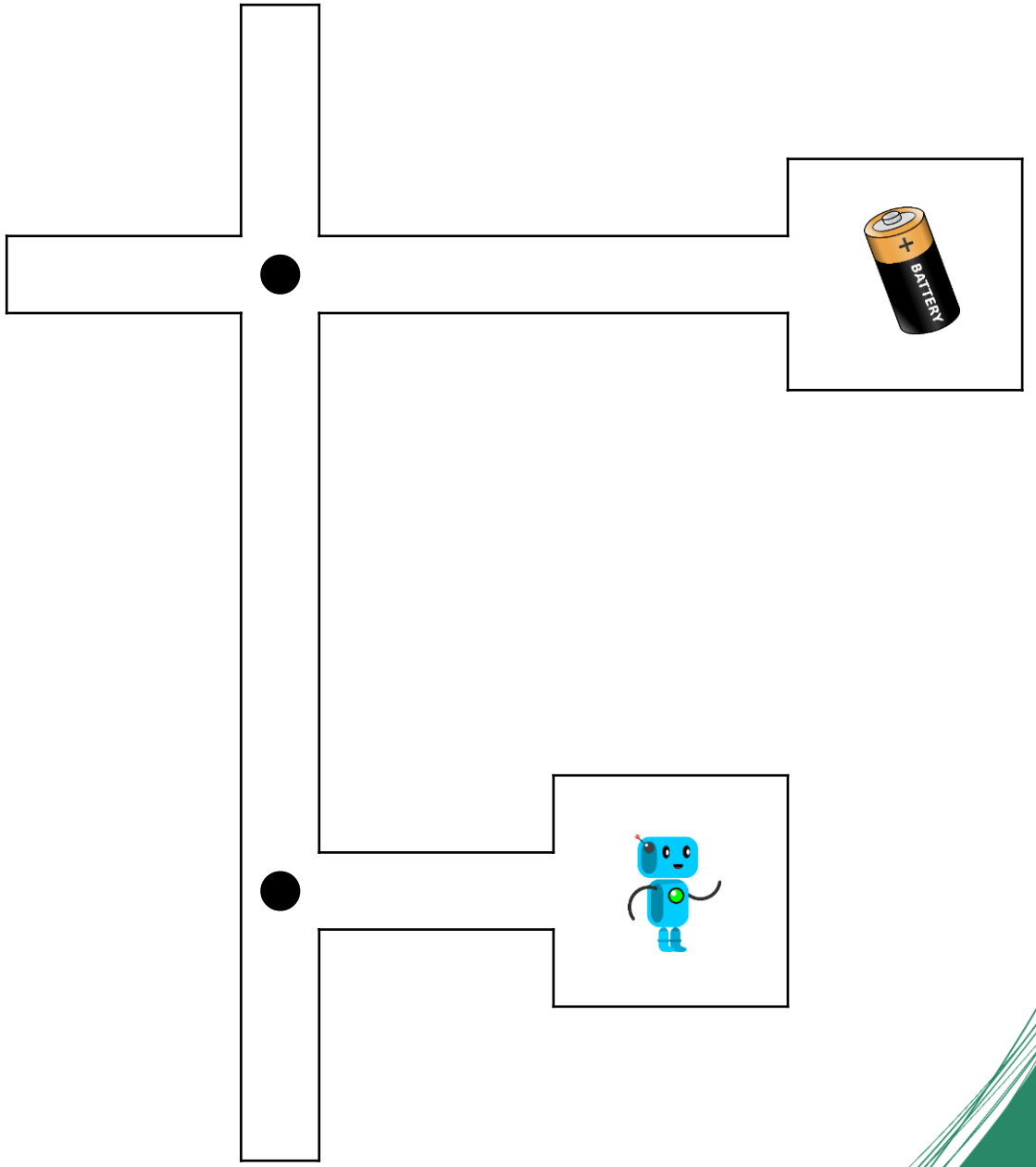
OESTERREICHISCHE COMPUTER GESELLSCHAFT ®
AUSTRIAN COMPUTER SOCIETY

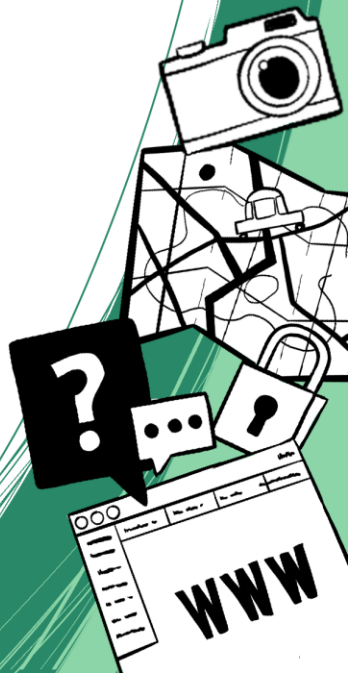
Interreg
Austria-Hungary 2014-2020
European Union – European Regional Development Fund

Alapszabályok

A játék

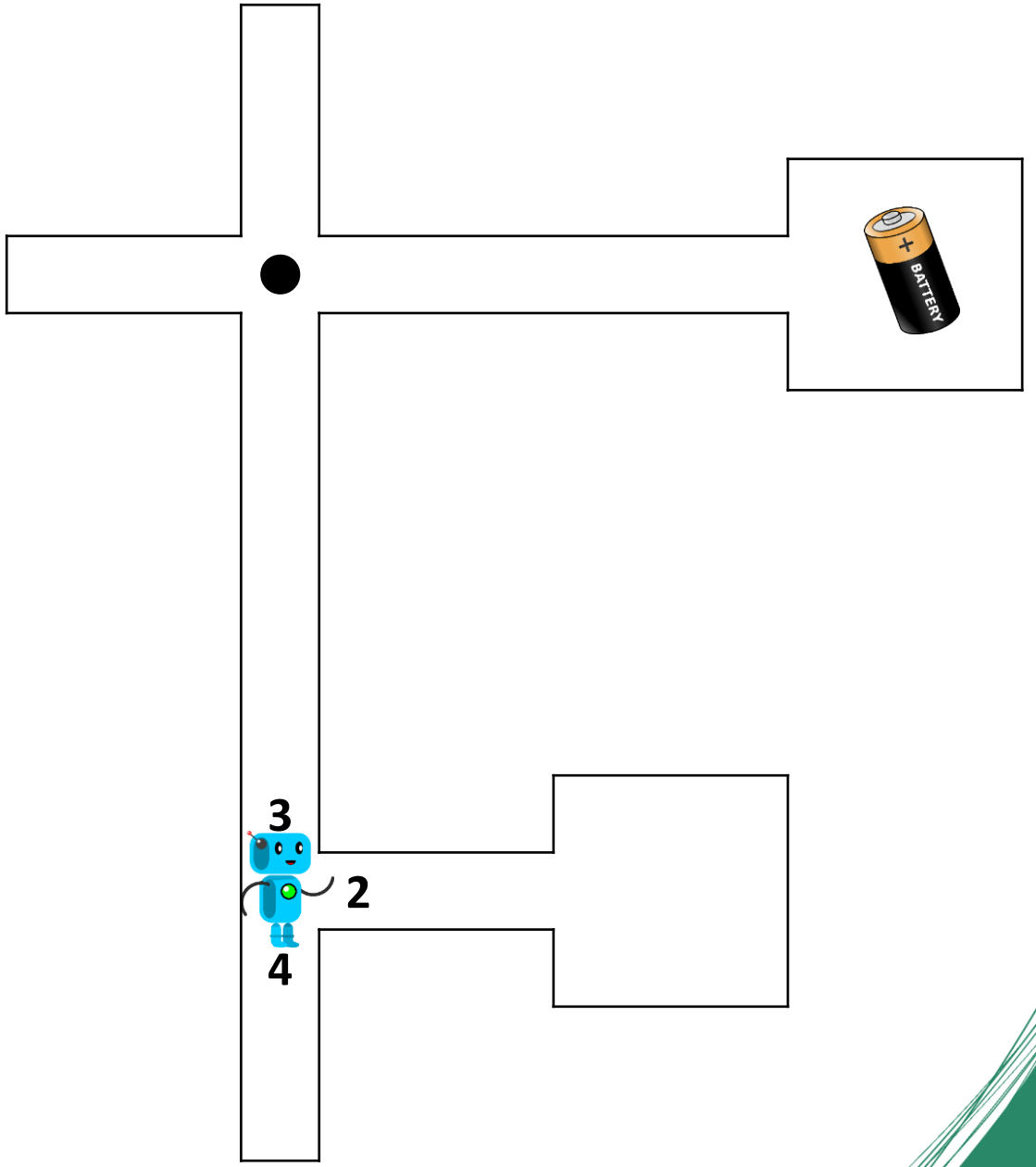
- **Ügynök**: robot
- **Cselekvések**: a következő útvonal kiválasztása
- **Állapot**: labirintus, robot pozíciója, Q-értékek
- **Jutalmak**: az útvonalakra írt számok
- **Cél**: megbízhatóan elérni az akkumulátort

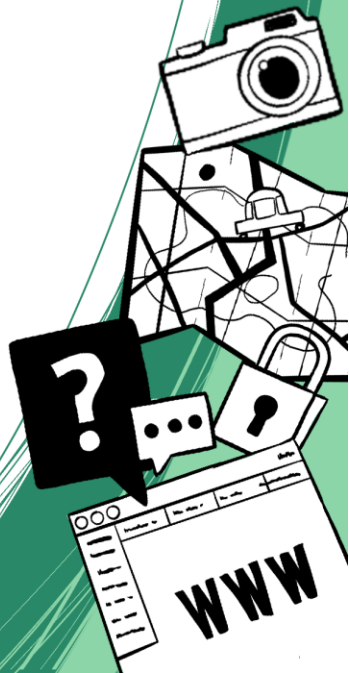




A játék

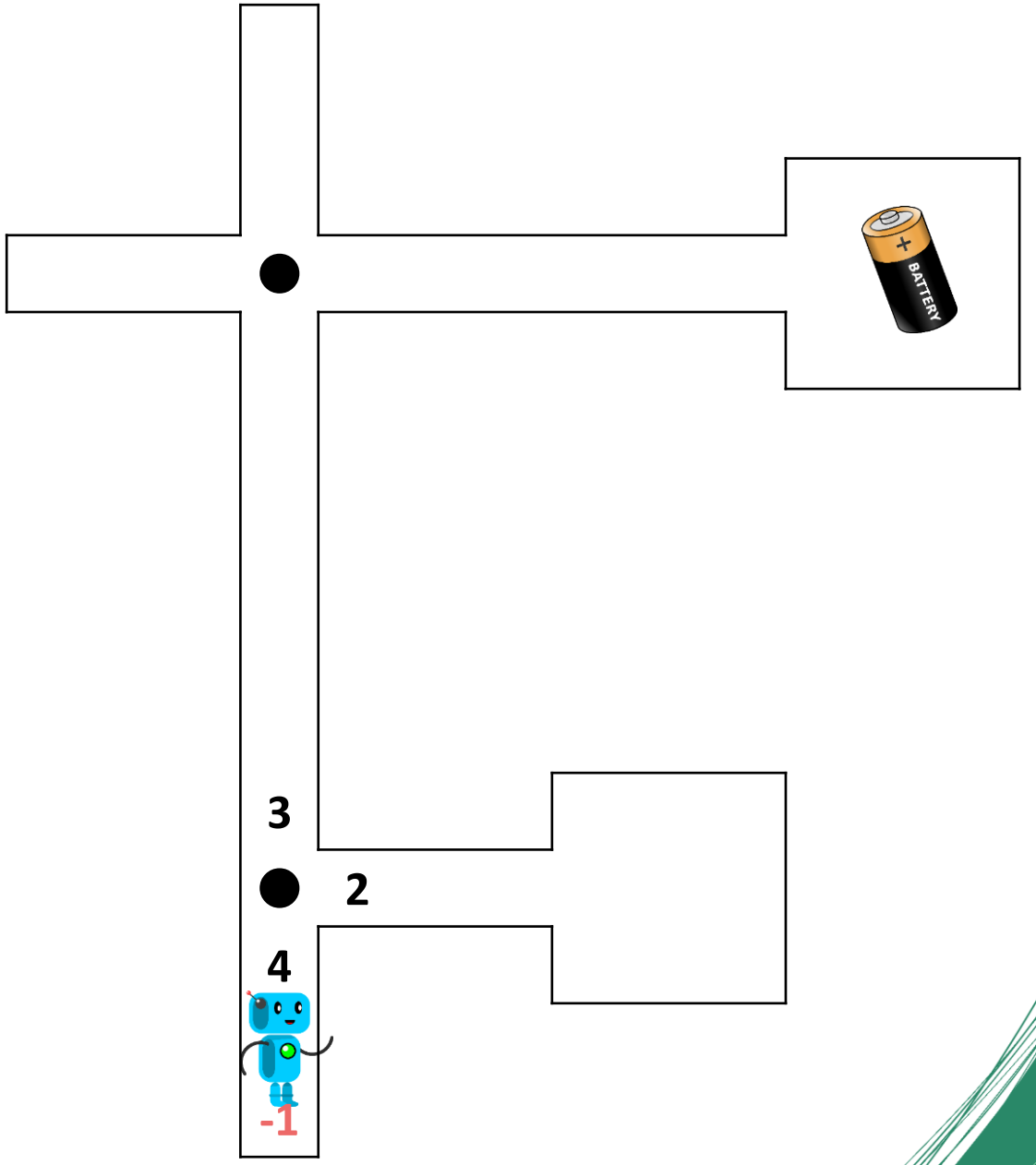
- Amikor a robot **új kereszteződéshez** ér, **minden útvonalra** **véletlen számot** ír (**dobókocka** segítségével)
- A robot ezután a **legmagasabb számmal** rendelkező utat választja

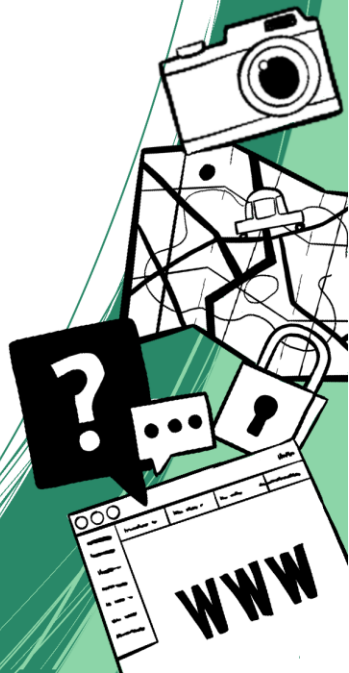




A játék

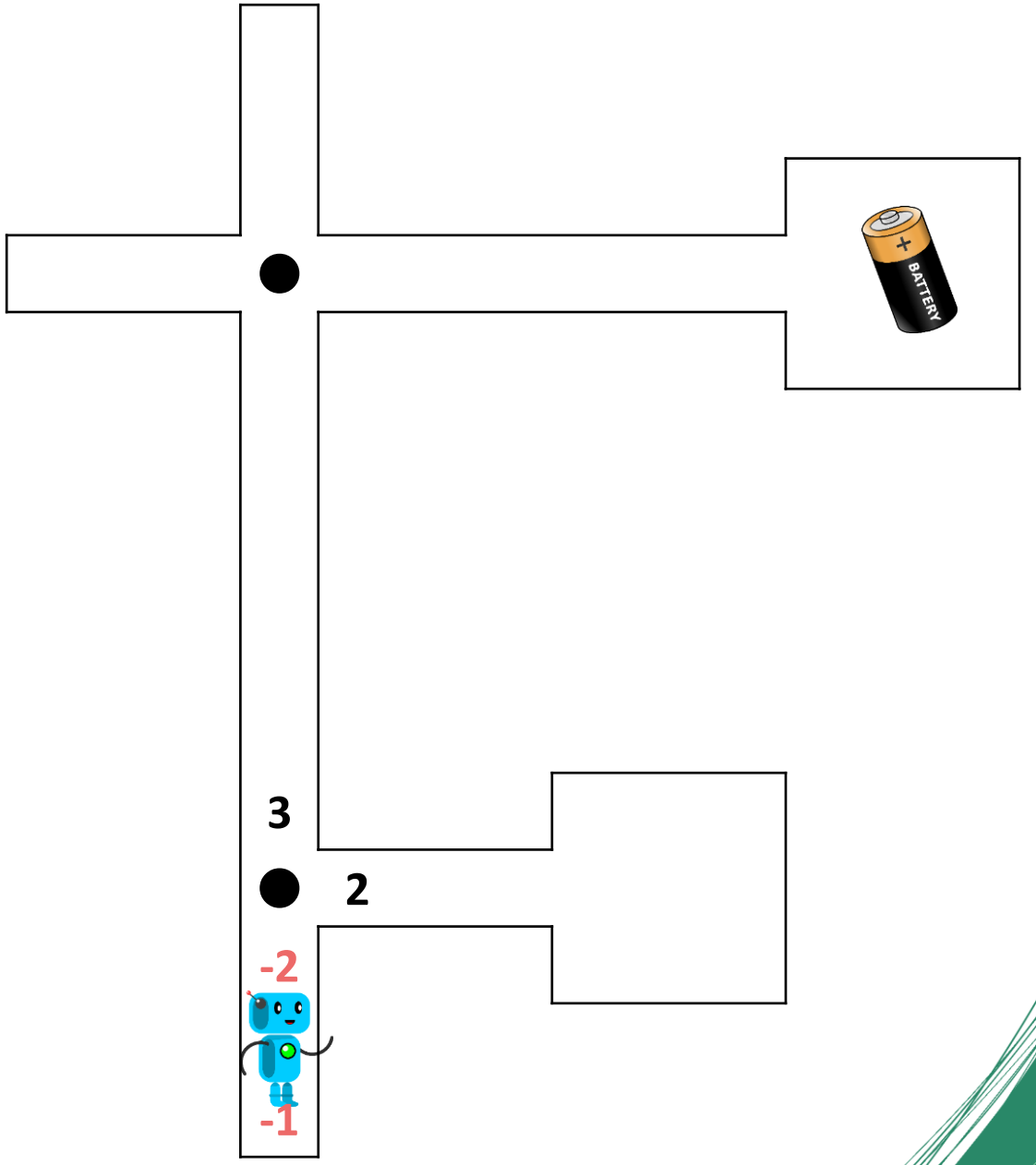
- Amikor a robot **zsákutcába** ér, **–1**-et ír

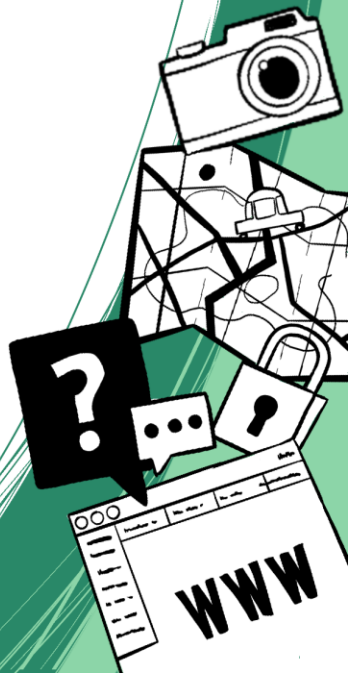




A játék

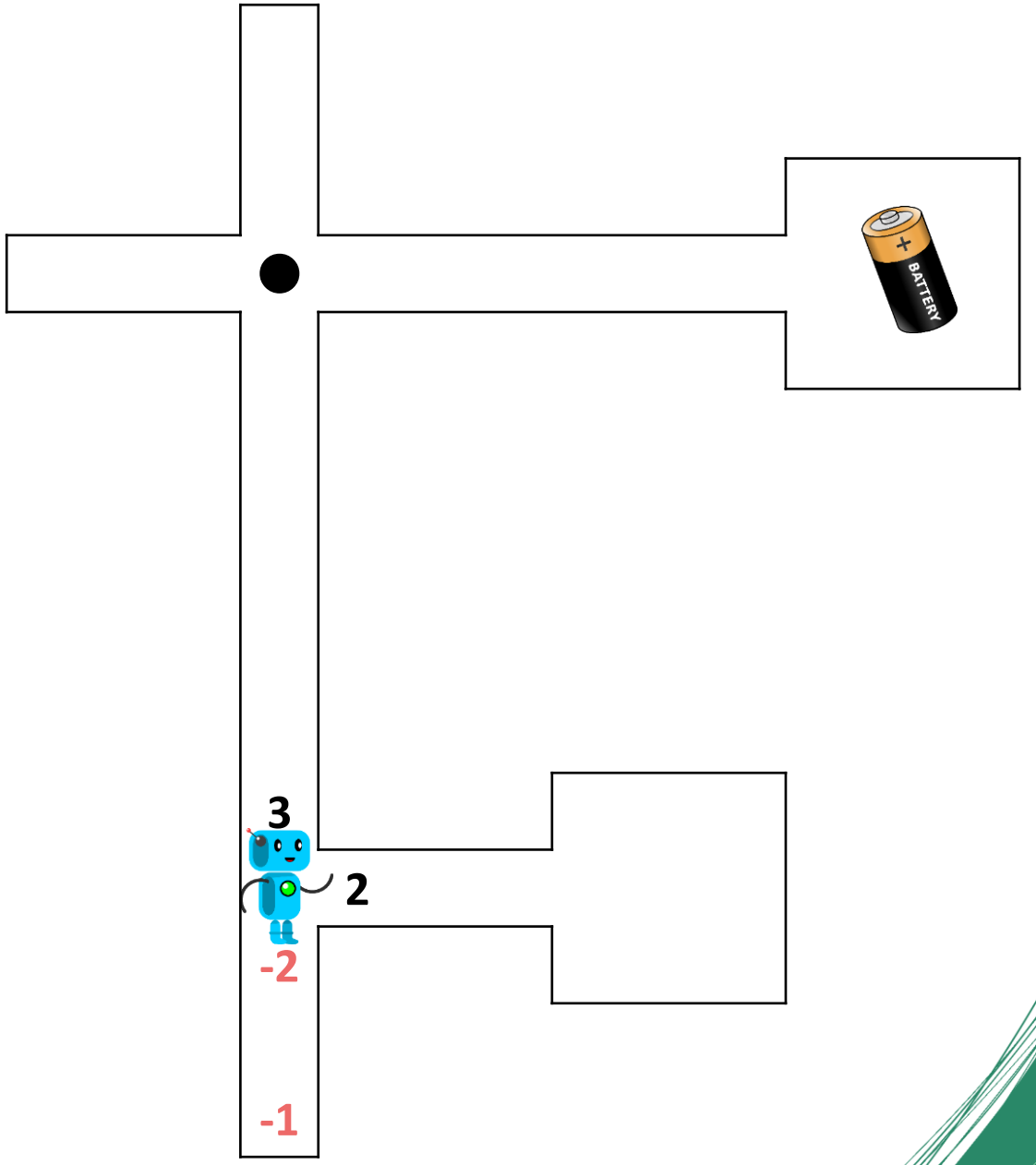
- Amikor a robot **zsákutcába** ér, **-1**-et ír
- Ezután a robot **frissíti** annak az **útvonalnak a számát, ahonnan jött**, az új **legmagasabb érték** (-1) mínusz egyre (-1-1=-2).





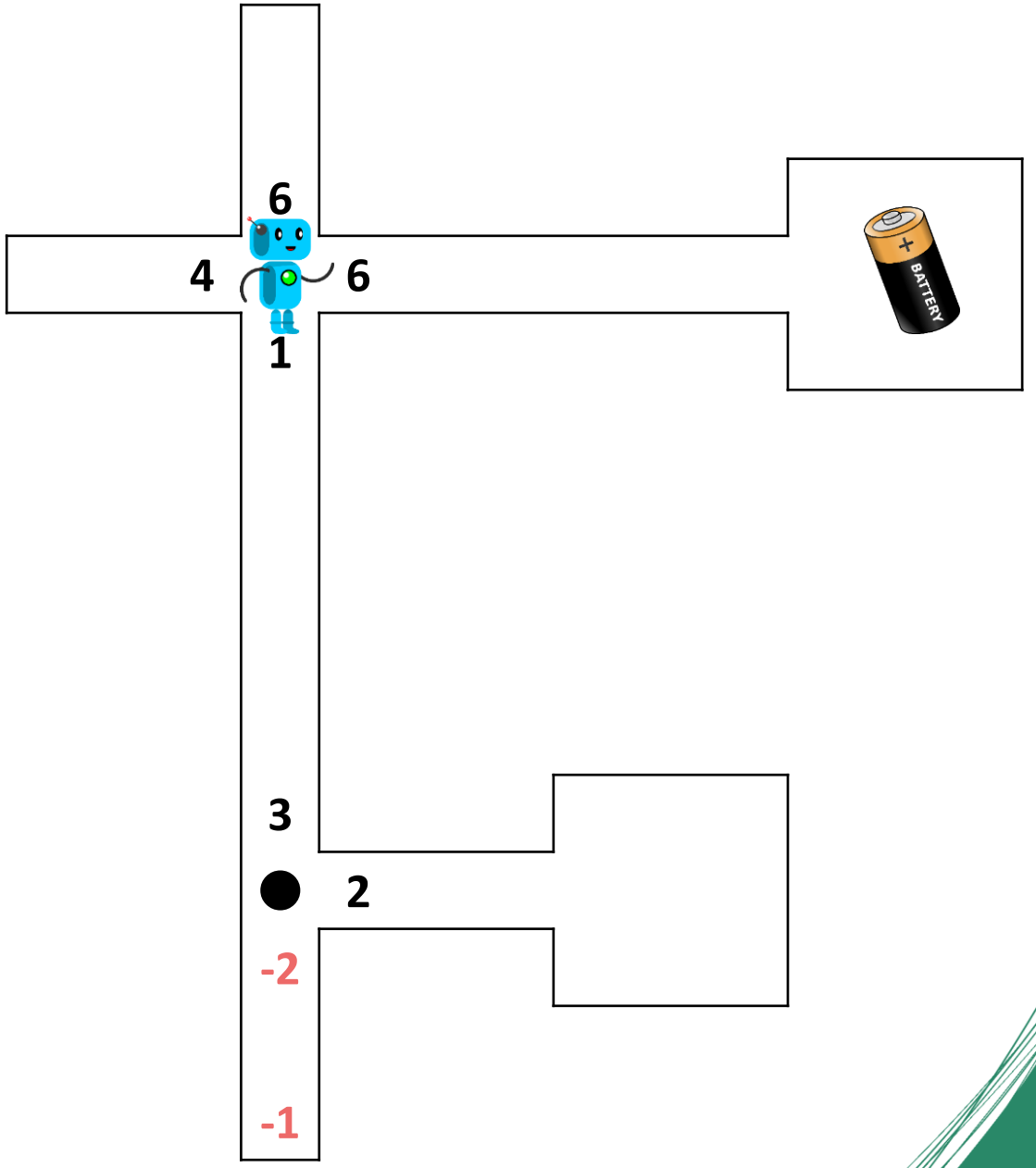
A játék

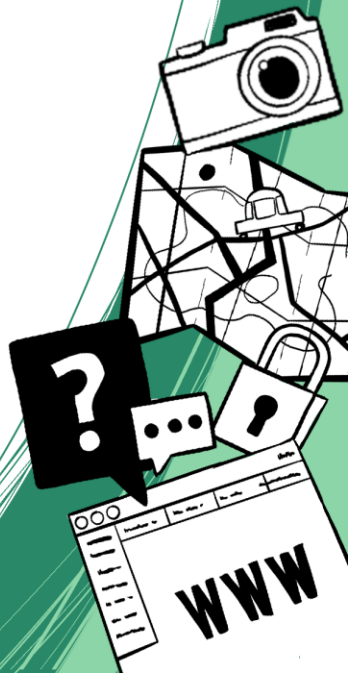
- **Zsákutca** esetén a robot visszatér az előző kereszteződéshez, és a legnagyobb számot választva folytatja



A játék

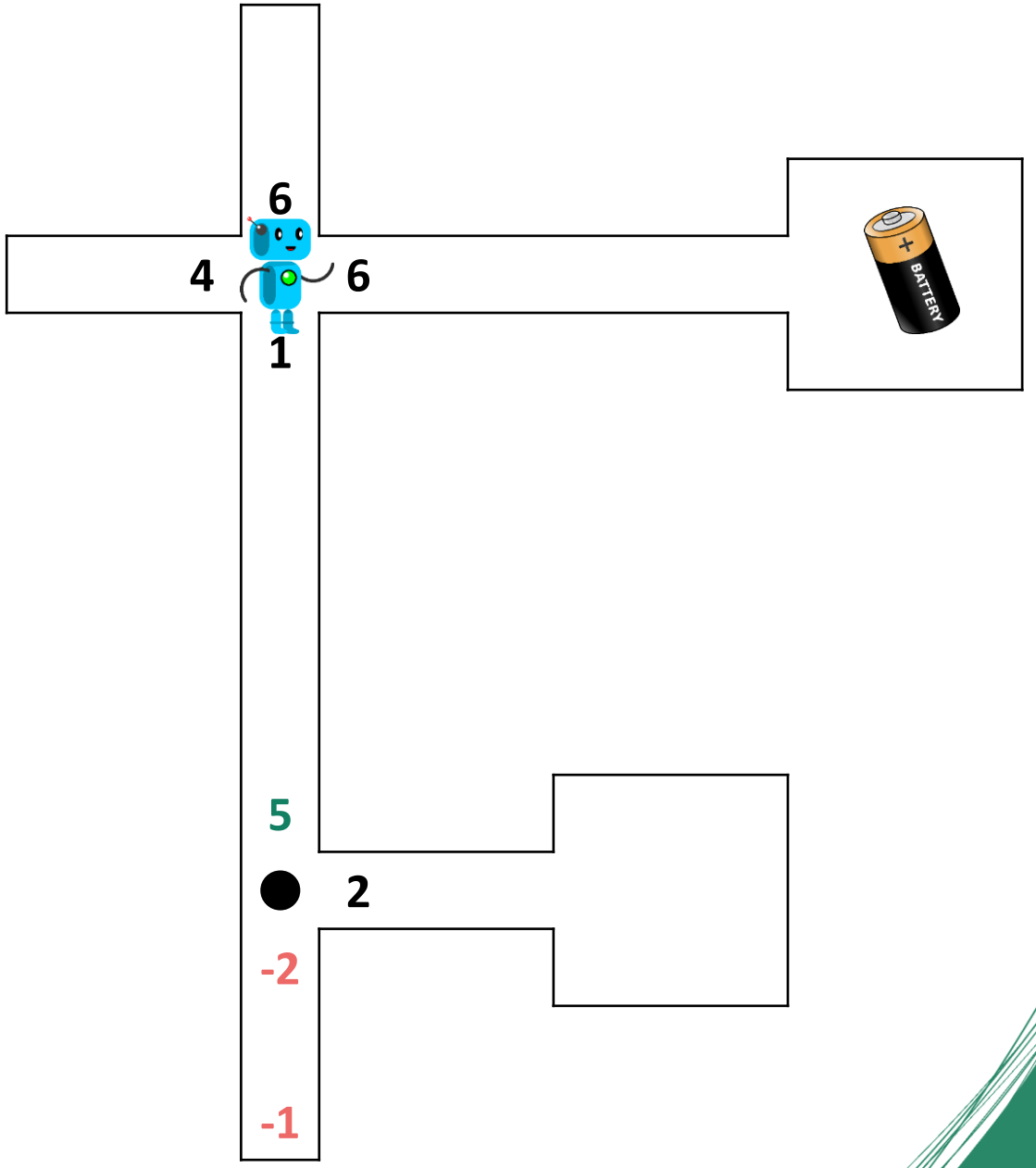
- **Új kereszteződés**-> véletlen számok

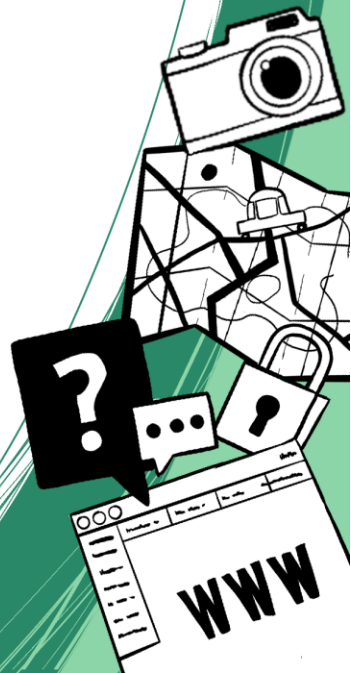




A játék

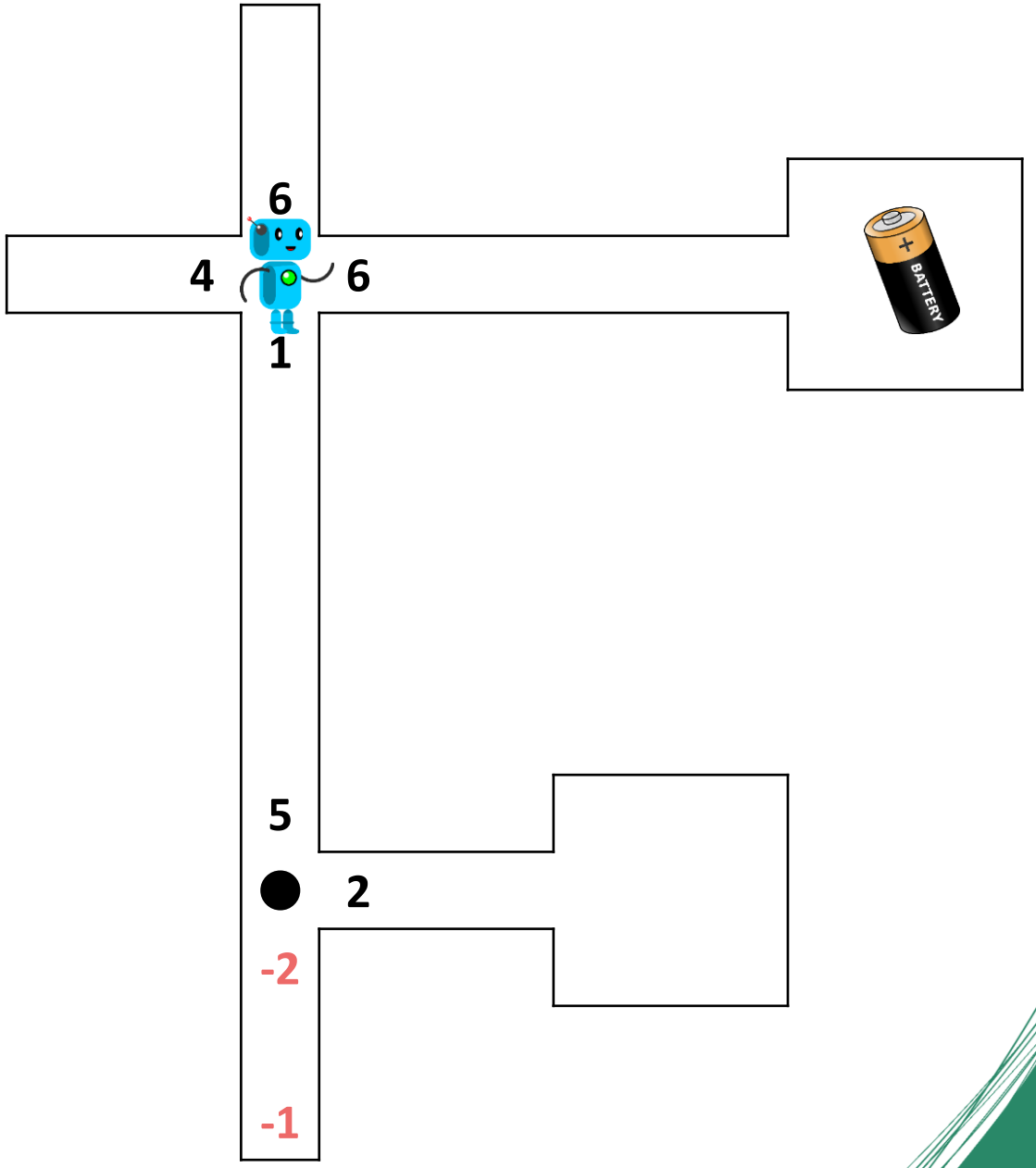
- **Új kereszteződés**-> véletlen számok
- Az útvonal **frissítése** az új legmagasabb (6) mínusz egyre (6-1=5)

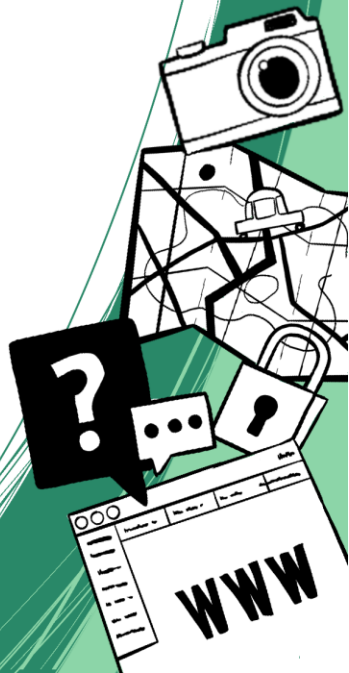




A játék

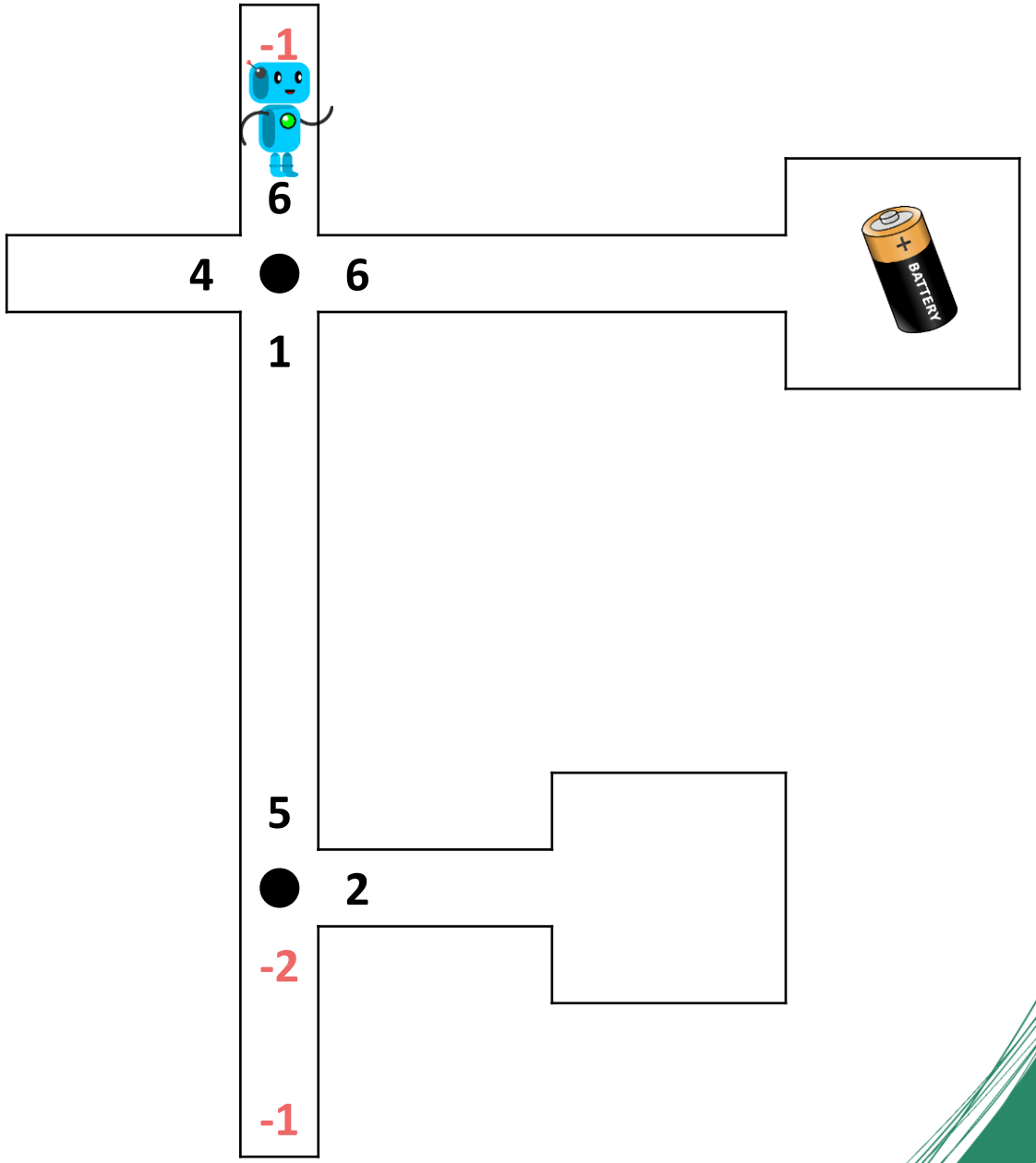
- **Új kereszteződés**-> véletlen számok
- Az útvonal **frissítése** az új legmagasabb (6) mínusz egyre (6-1=5)
- Ha **kettő vagy több** legnagyobb **szám** van, a robot **véletlenszerűen** választ egyet (ebben az esetben: fel).

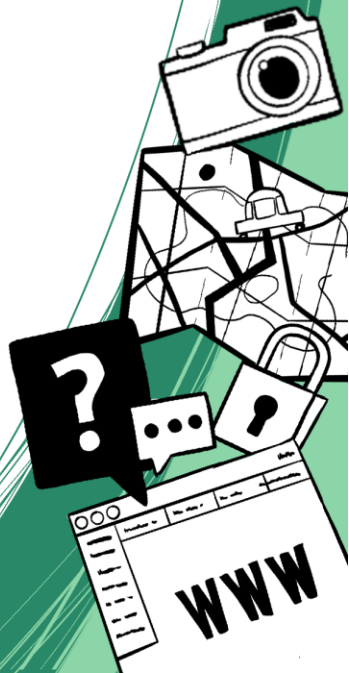




A játék

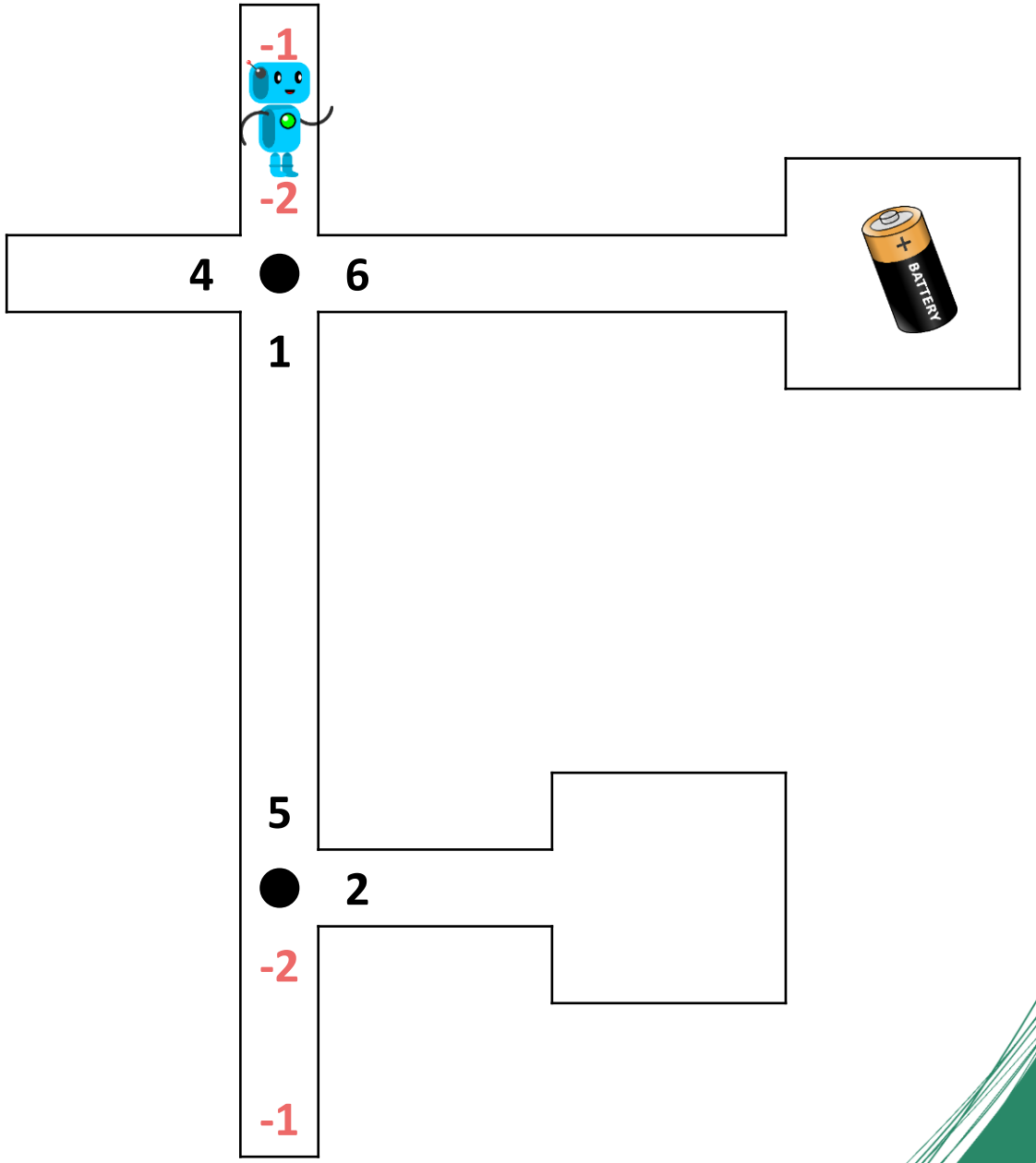
- A robot addig **folytatja**, amíg el nem éri a **célját**

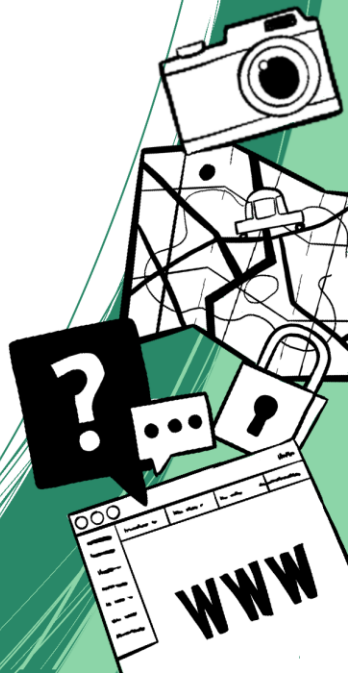




A játék

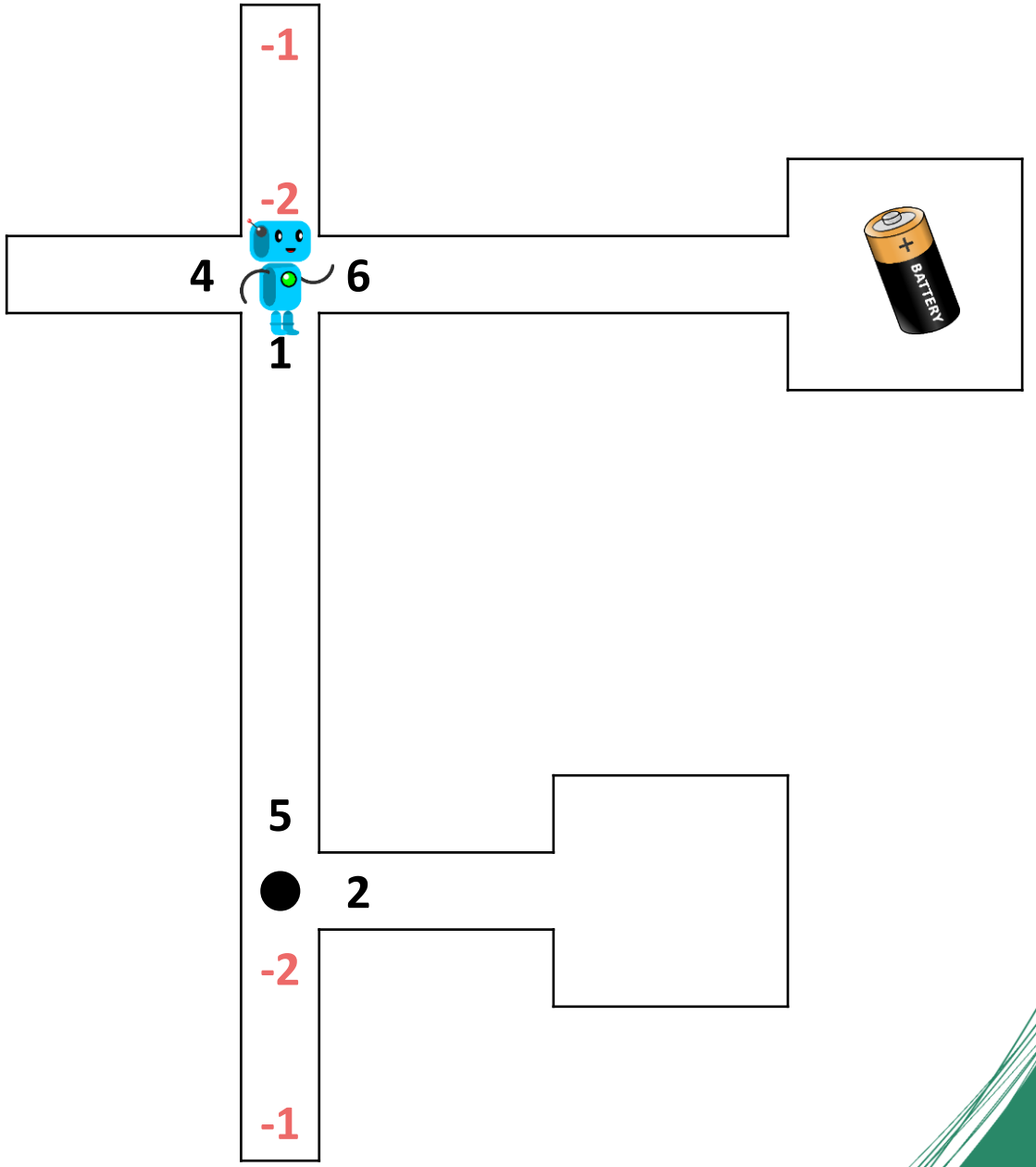
- A robot addig **folytatja**, amíg el nem éri a **célját**





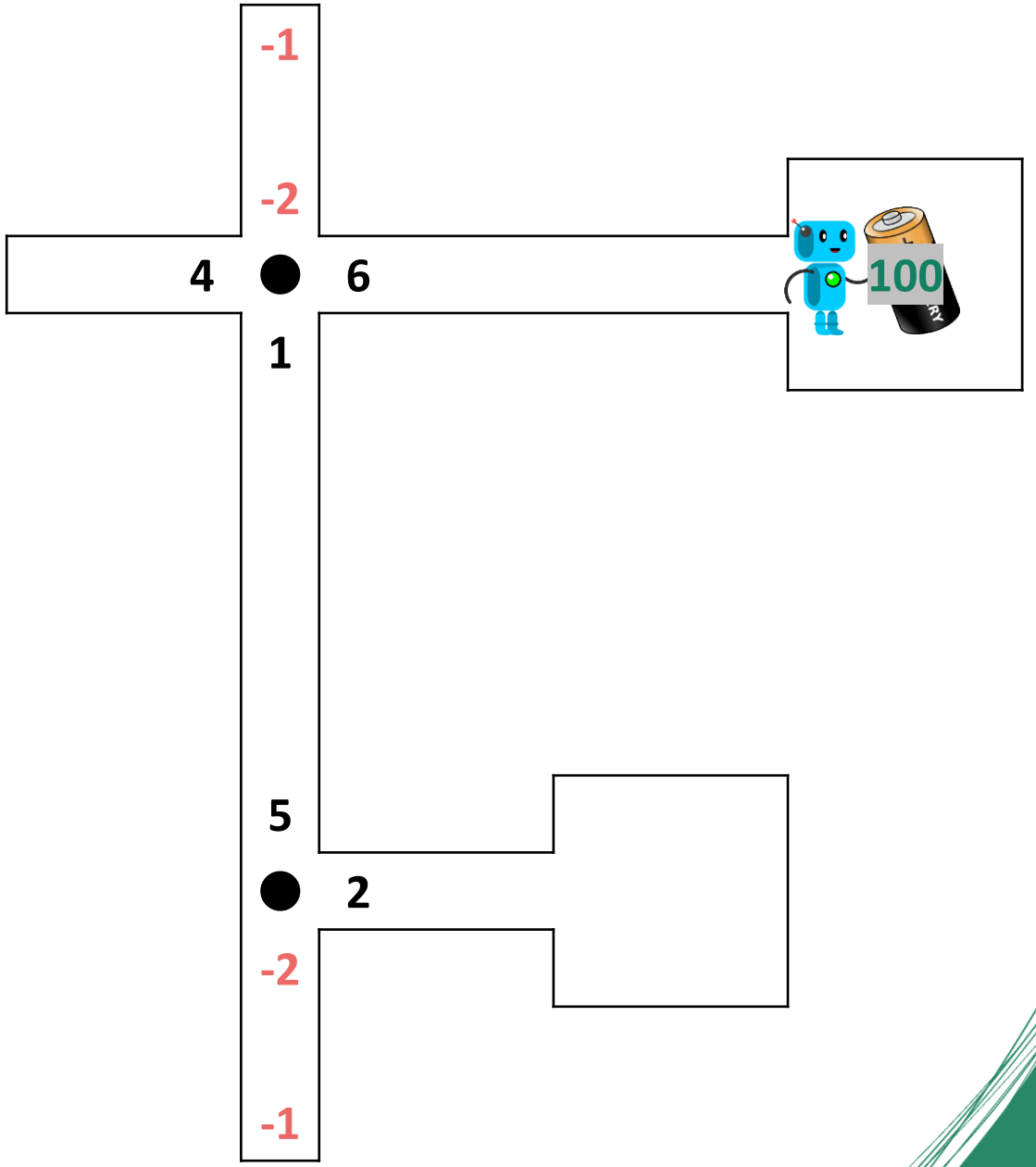
A játék

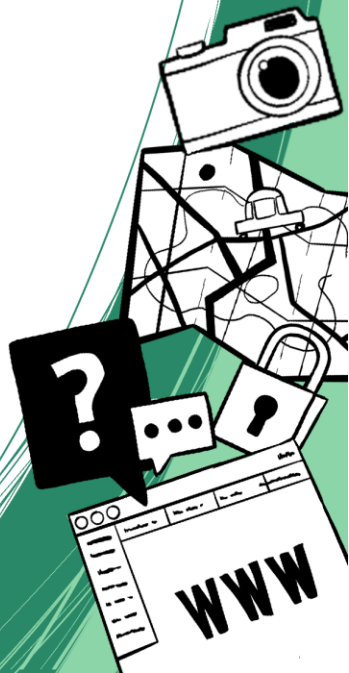
- A robot addig **folytatja**, amíg el nem éri a **célját**



A játék

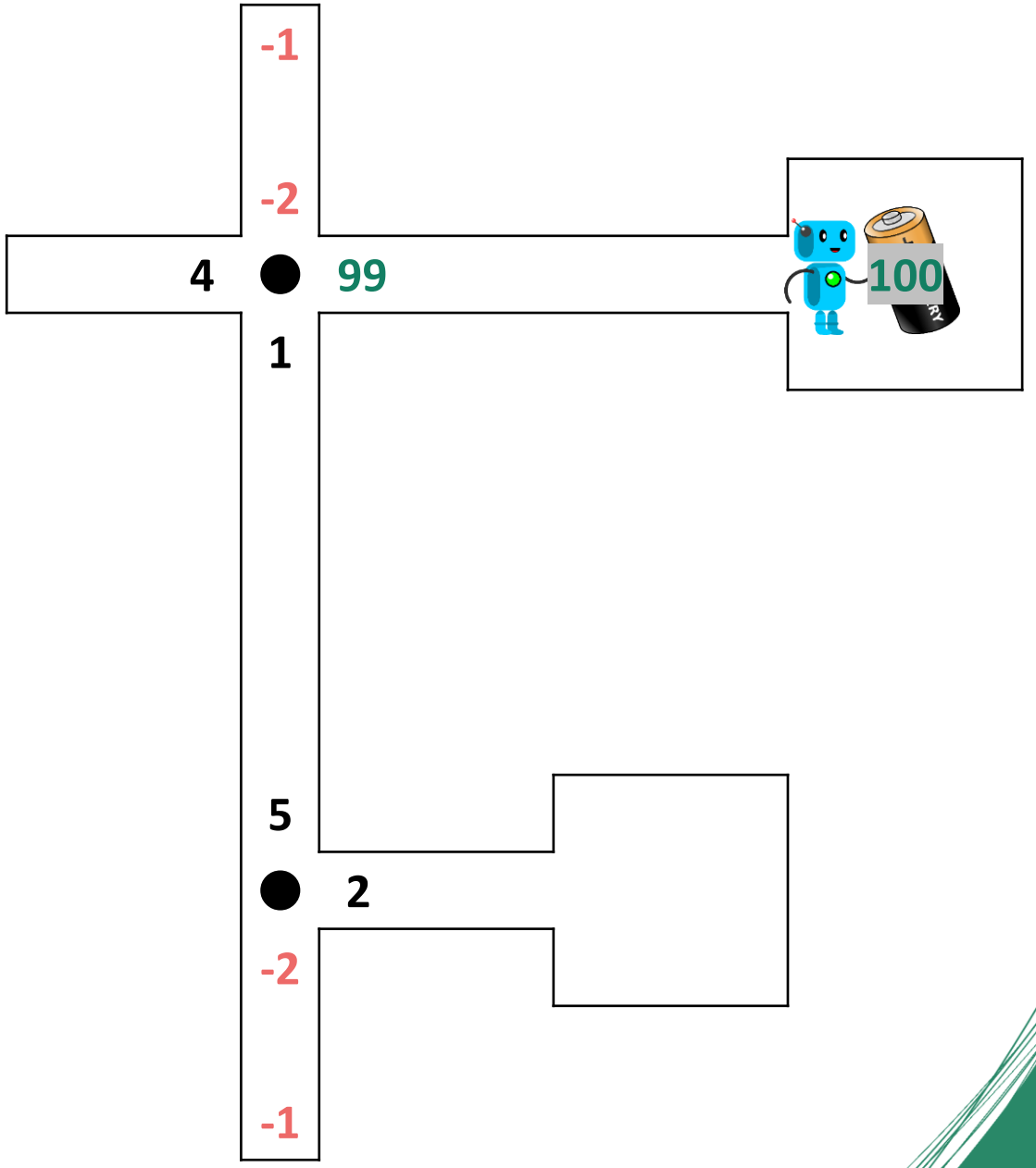
- A cél ebben a labirintusban **100-as** értékű





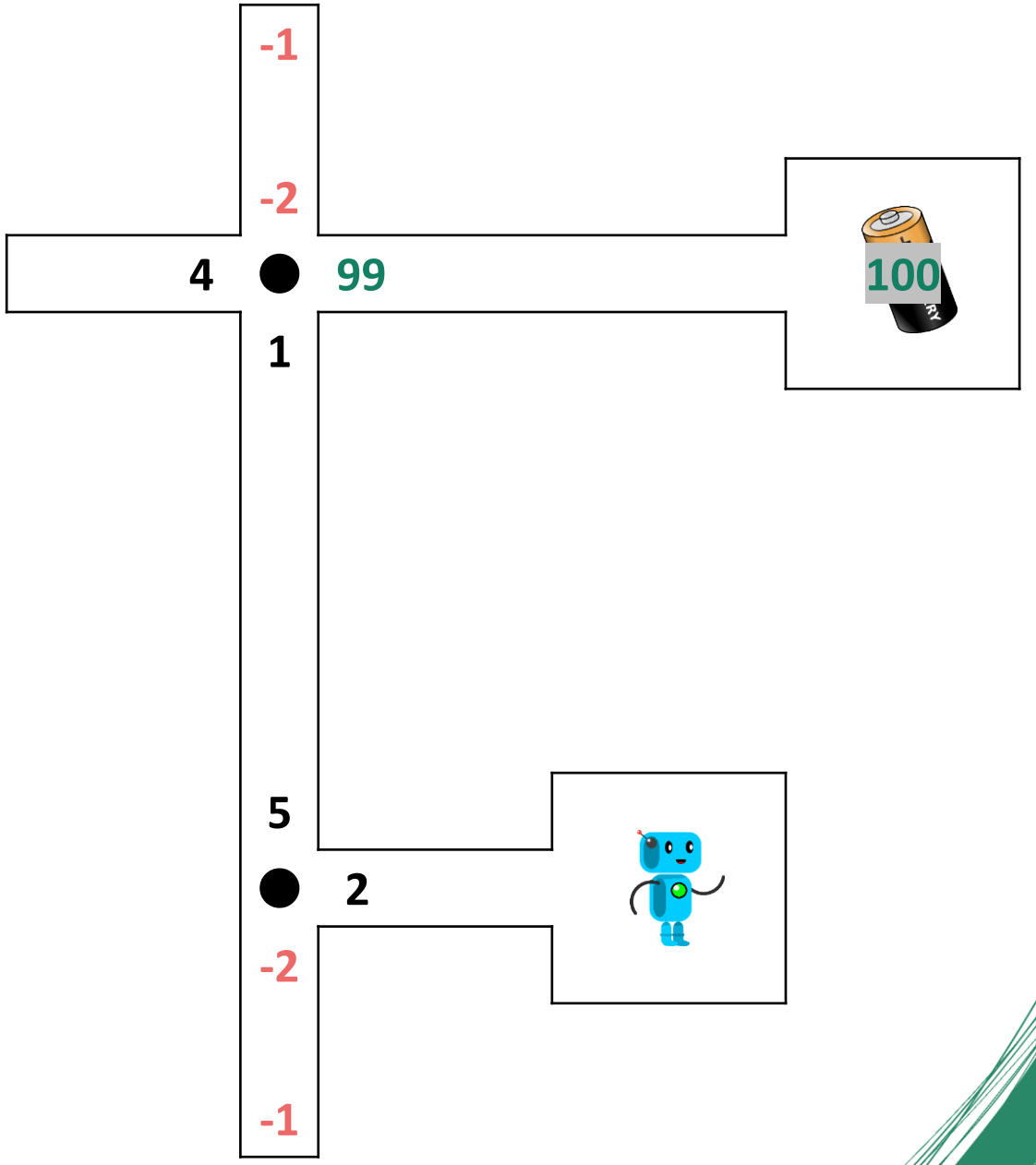
A játék

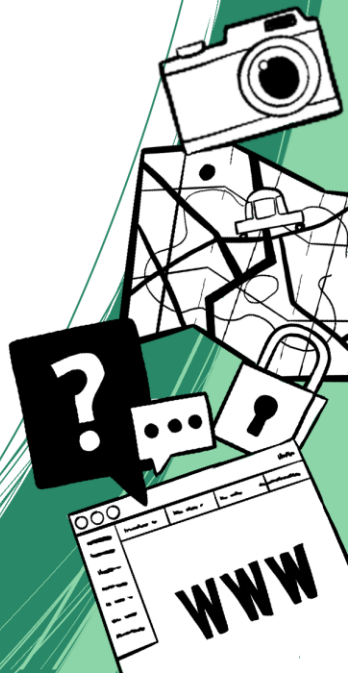
- A cél ebben a labirintusban **100-as** értékű
- Ne felejtsük el **frissíteni** az útvonalat!



A játék

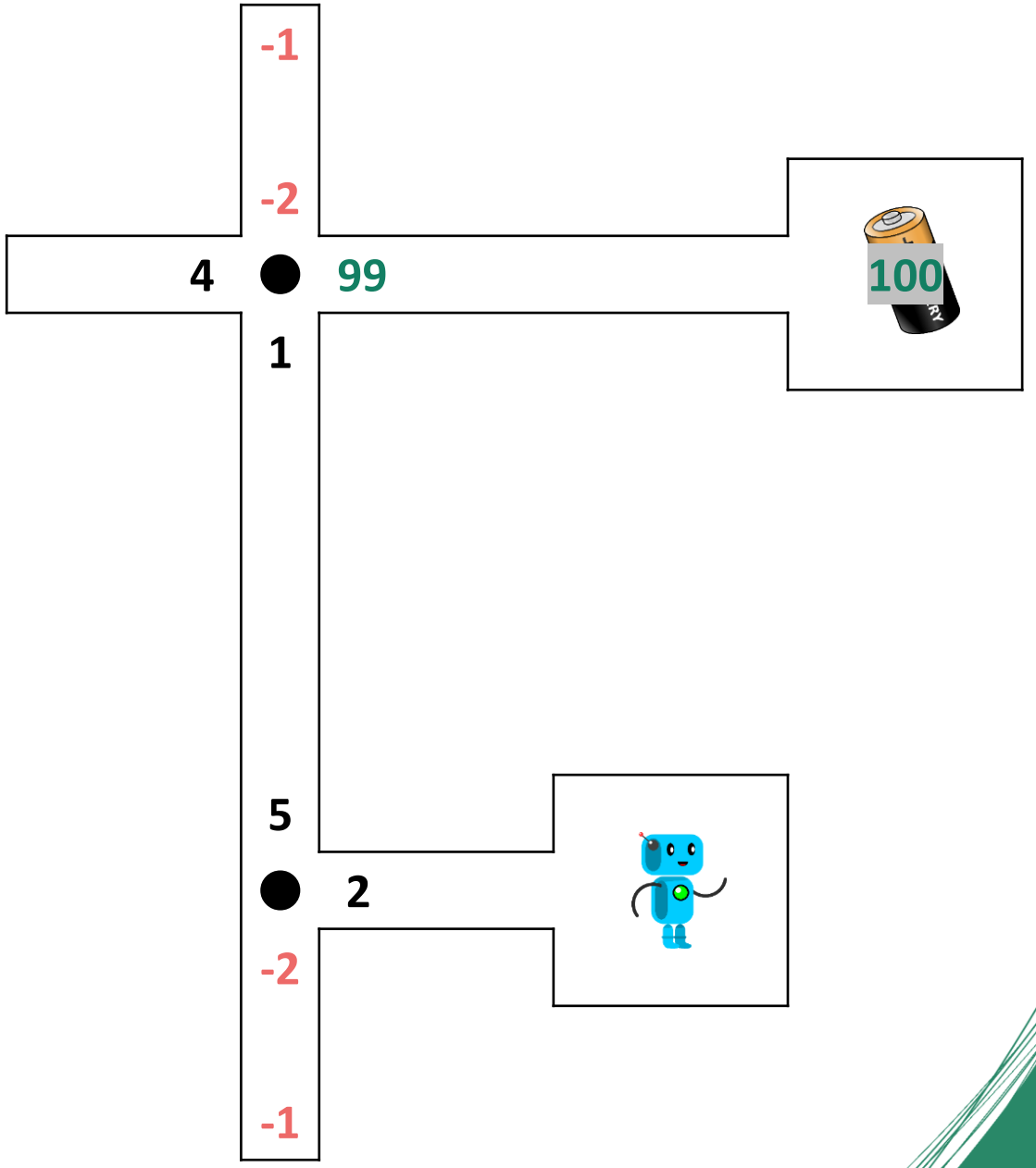
- A cél ebben a labirintusban **100-as** értékű
- Ne felejtsük el **frissíteni** az útvonalat!
- Ezután a robot **visszatér** az **első helyiségbe**, és újabb kört kezd

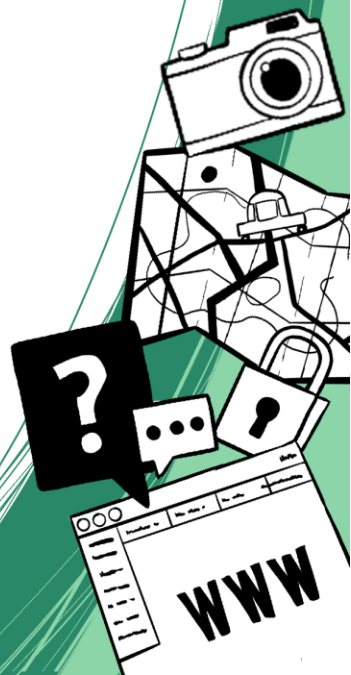




A játék

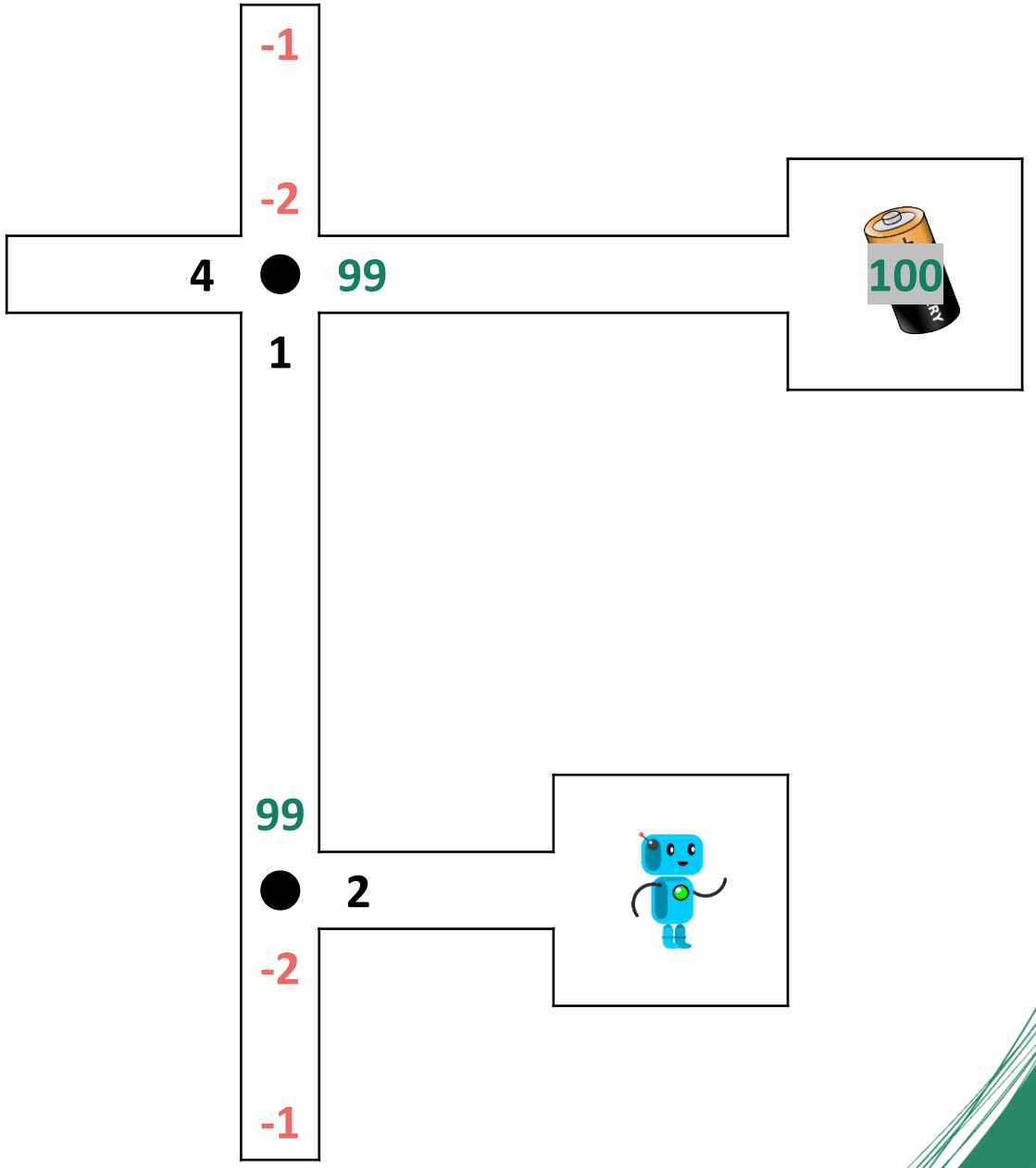
- **Folytassuk**, amíg a robot **nem tanul semmi újat** (számokat írni vagy javítani).

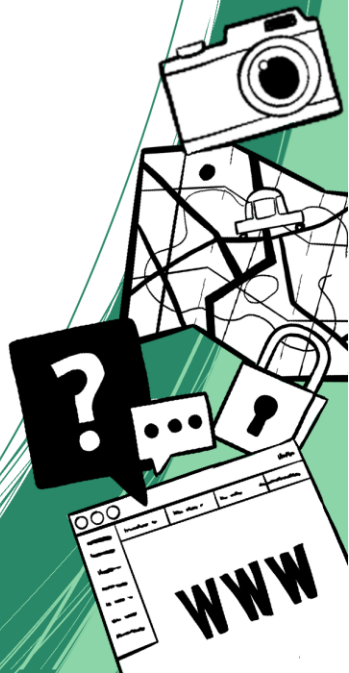




A játék

- Végül a robot **megtanult** egy megbízható **utat a céljához**





Felfedezés vs. kiaknázás

A probléma

- Egyes **cselekvéseknek** lehet **pozitív** hatása **azonnal**, de **negatív hatása hosszú távon**

A probléma

- Egyes **cselekvéseknek** lehet **pozitív** hatása **azonnal**, de **negatív hatása hosszú távon**
  - Pl. a robot elindul egy ösvényen, és folytatja az utat, de sok lépéssel később kiderül, hogy zsákutca

A probléma

- Egyes **cselekvéseknek** lehet **pozitív** hatása **azonnal**, de **negatív hatása hosszú távon**
  - Pl. a robot elindul egy ösvényen, és folytatja az utat, de sok lépéssel később kiderül, hogy zsákutca.
- Néhány **cselekvésnek** csak **kis pozitív** hatása van **azonnal**, de **nagy pozitív hatása hosszú távon**

A probléma

- Egyes **cselekvéseknek** lehet **pozitív** hatása **azonnal**, de **negatív hatása hosszú távon**
  - Pl. a robot elindul egy ösvényen, és folytatja az utat, de sok lépéssel később kiderül, hogy zsákutca.
- Néhány **cselekvésnek** csak **kis pozitív** hatása van **azonnal**, de **nagy pozitív hatása hosszú távon**
  - Pl. a robot talál egy értékes akkumulátort, de a következő útvonal mentén lenne egy még értékesebb.

A megoldás

- Találjuk meg az **egyensúlyt** a már megtanult információk **kiaknázása** és az új lehetőségek **felfedezése** között

A megoldás

- Találjuk meg az **egyensúlyt** a már megtanult információk **kiaknázása** és az új lehetőségek **felfedezése** között
  - Pl. ahelyett, hogy mindig a **legmagasabb számmal** rendelkező pályát választaná, a robot **25%-os eséllyel választhatna egy véletlenszerű pályát**

A megoldás

- Találjuk meg az **egyensúlyt** a már megtanult információk **kiaknázása** és az új lehetőségek **felfedezése** között
  - Pl. ahelyett, hogy mindig a **legmagasabb számmal** rendelkező pályát választaná, a robot **25%-os eséllyel választhatna egy véletlenszerű pályát**
  - Ez a **felfedezési arány** lehet dinamikus is, így **kezdetben magas**, de **csökkenő**, ahogy a robot egyre jobban megismeri a környezetet.